Progress in using and understanding jets

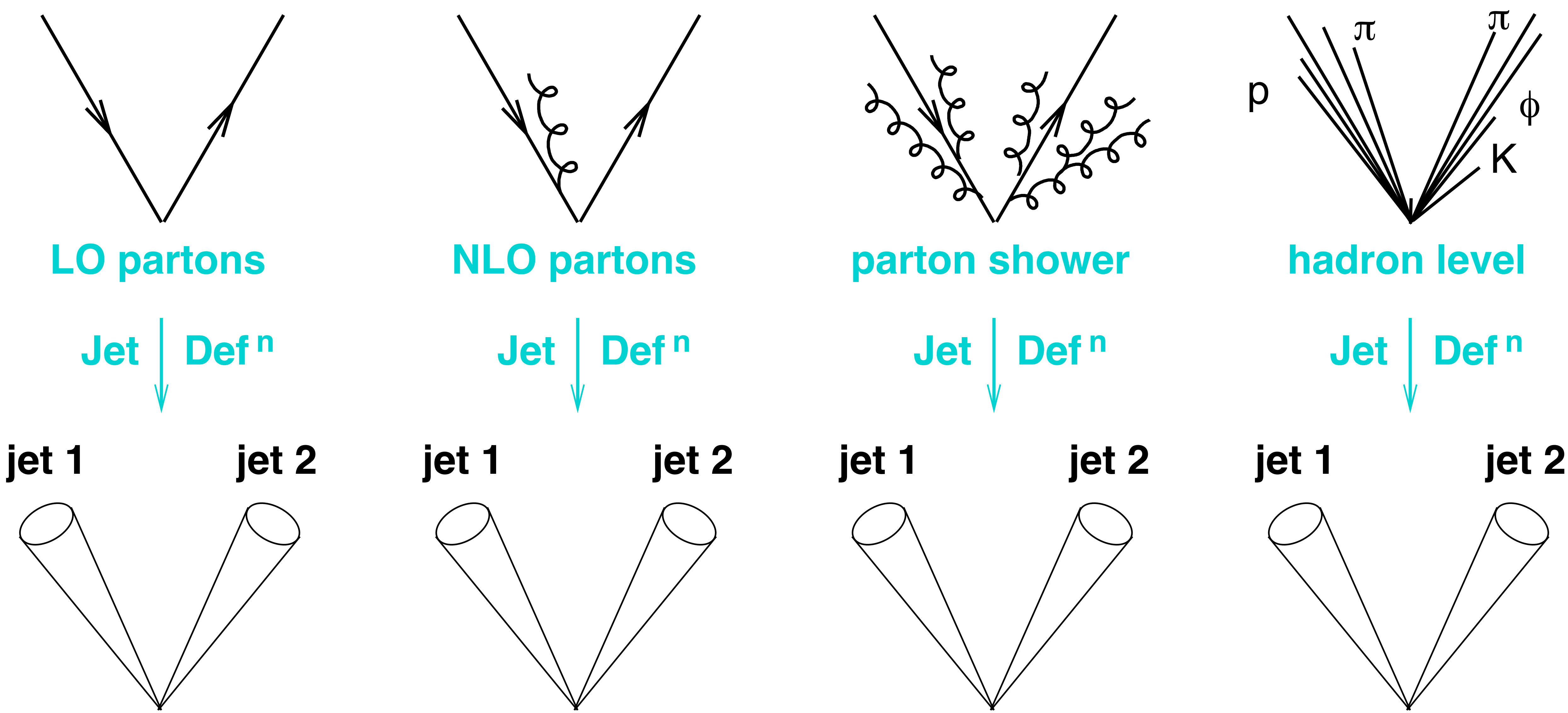
Gavin Salam (CERN)

The First Three Years of the LHC
Mainz Institute for Theoretical Physics
Johannes-Gutenberg-Universität
18–22 March 2013

60-70% of recent ATLAS and CMS papers use jets in their analyses

[Higgs, BSM searches, top physics, SM studies, heavy-ions, ...]

Jets project away "useless" information



Projection to jets should be resilient to QCD effects

Most LHC jet uses fall under the (Tevatron-like) category

"a jet is basically a parton"

e.g. from a heavy-object decay, ISR, etc.

If radiation is modelled correctly in your Monte Carlo
[or NLO/NNLO/resummation – see next talks]
most experimenters don't even need to think (much) about jets.

Most LHC jet uses fall under the (Tevatron-like) category

"a jet is basically a parton"

e.g. from a heavy-object decay, ISR, etc.

If radiation is modelled correctly in your Monte Carlo
[or NLO/NNLO/resummation – see next talks]
most experimenters don't even need to think (much) about jets.

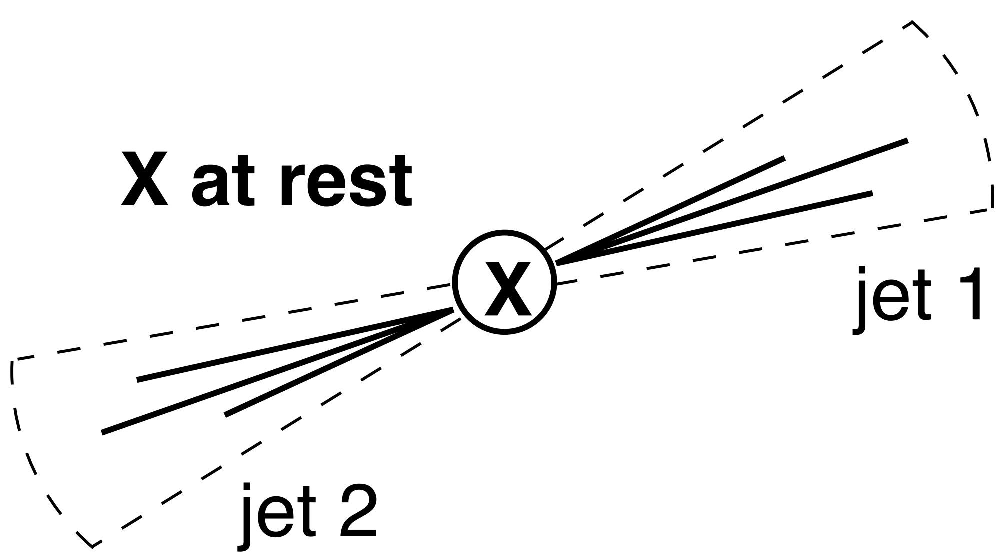
But with new physics proving elusive, we start to need to push analyses to their boundary, e.g.

Enhance sensitivity to small signal/background
Explore very highest p_t's
Learn how to handle complex final states

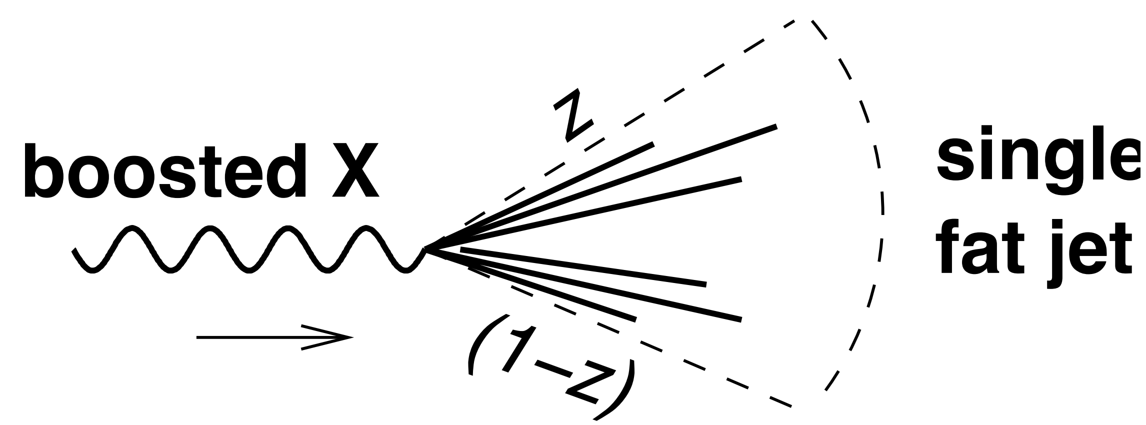
➜ for that, you need advanced jet techniques

Boosted hadronic decays

Normal analyses: two quarks from $X \to q\bar{q}$ reconstructed as two jets



High-p_t regime: EW object X is boosted, decay is collimated, $q\bar{q}$ both in same jet

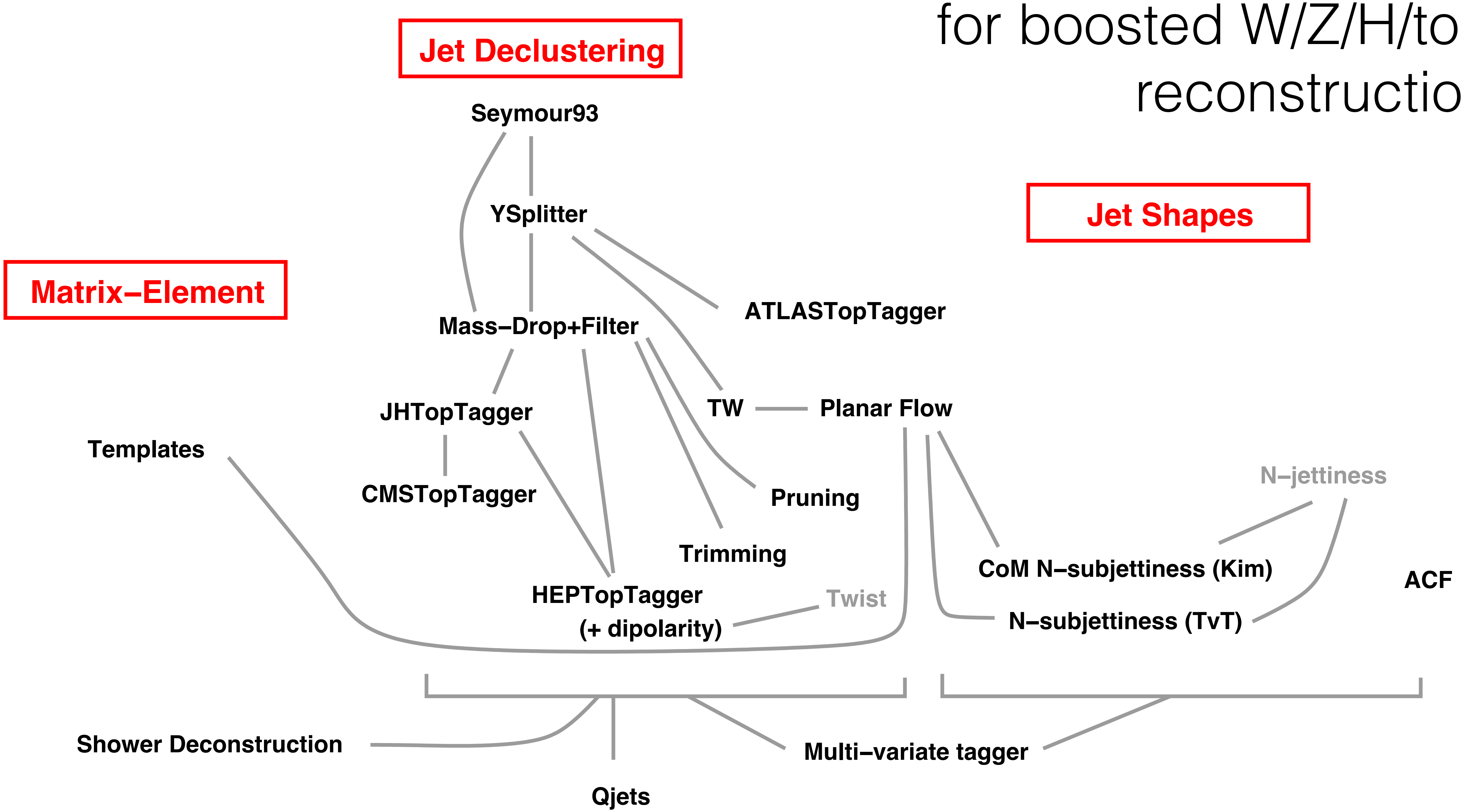


Happens for $p_t \gtrsim 2m/R$

$p_t \gtrsim 320$ GeV for $m = m_W$, $R = 0.5$

Very active research field

Some of the tools developed for boosted W/Z/H/top reconstruction



apologies for omitted taggers, arguable links, etc.

Extensive experimental work

ATLAS Public Results

- Large-R, groomed jets with pile-up
- Large-R jets with substructure
- Quark/gluon jets (see also this link)
- Jet substructure at LHC7
- Jet properties for boosted searches

Resonance searches

- Boosted top (hadronic)
- Boosted top (semileptonic)
- Three-jet resonance (gluino RPV)
- Two-jet resonance (sgluon)

CMS Public Results

- Jet substructure in CMS
- Subjet multiplicity
- Jet mass and grooming

Resonance searches:

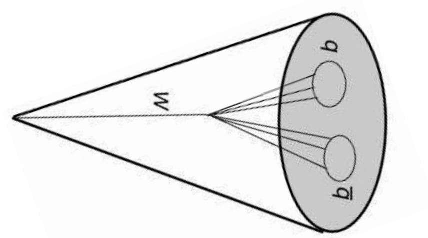
- Boosted top (hadronic)
- Boosted top (semileptonic)
- Boosted W/Z

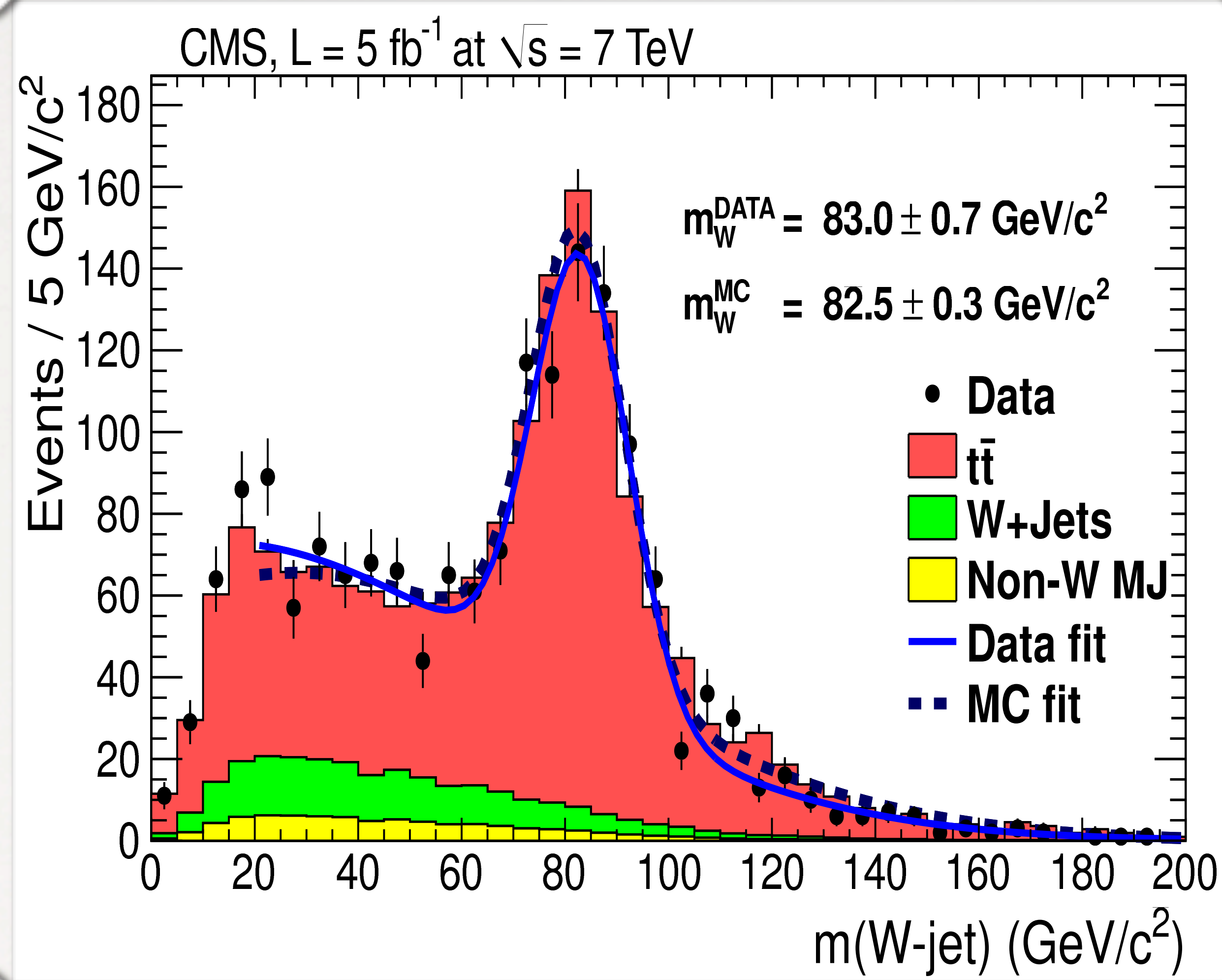
From a list compiled for a recent workshop at Perimeter Institute

Many more analyses in the pipeline

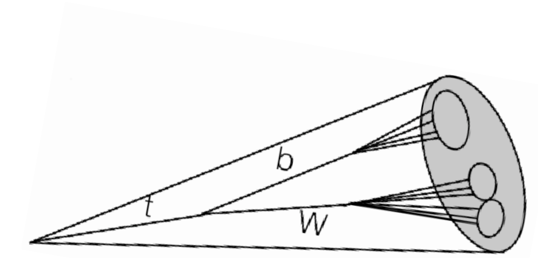
Seeing W's and tops in a single jet

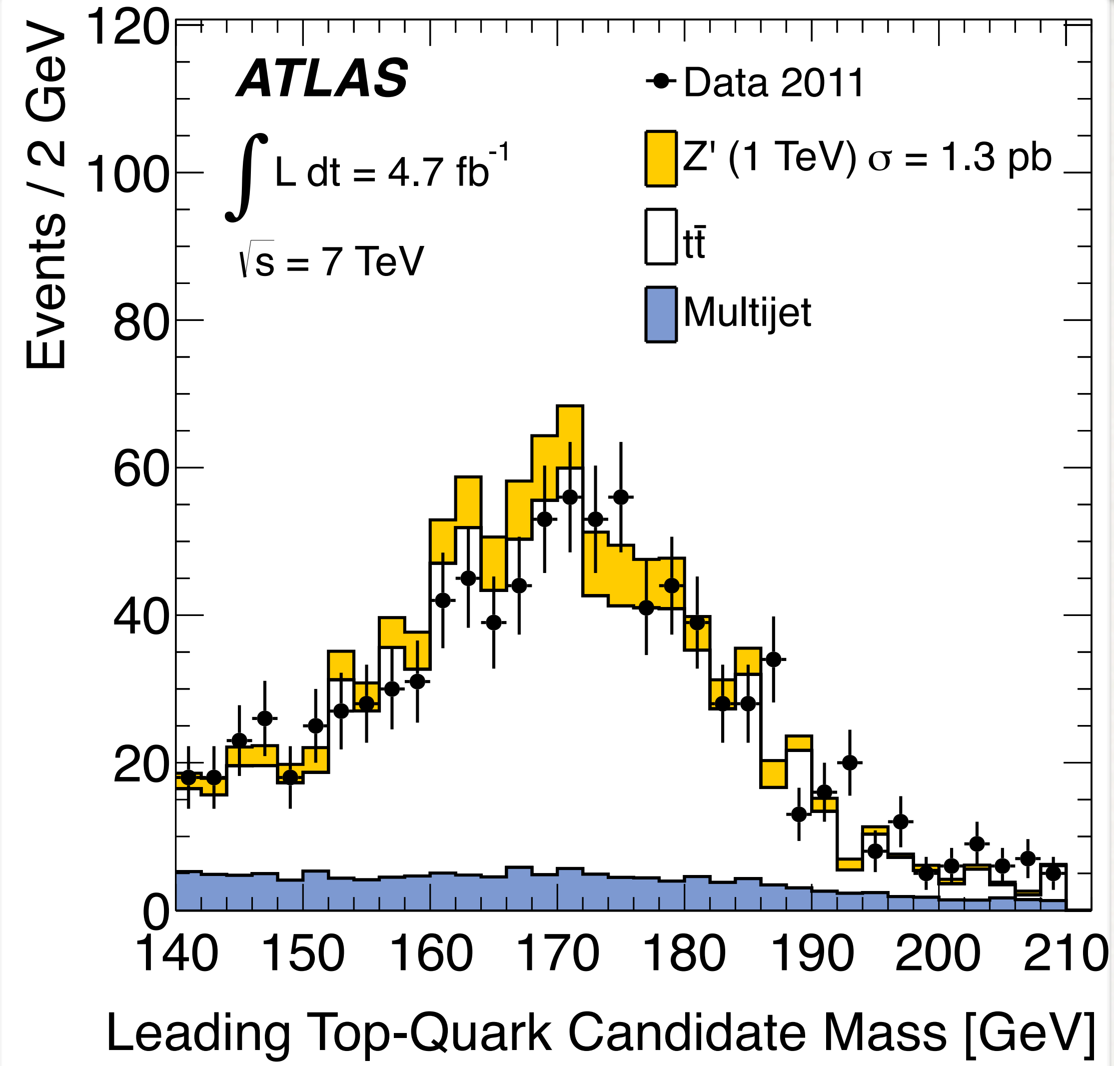
W's in a single jet





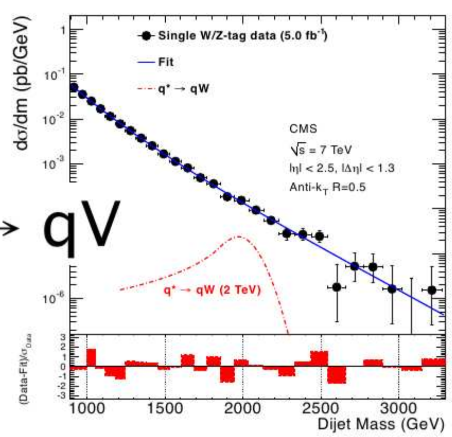
tops in a single jet

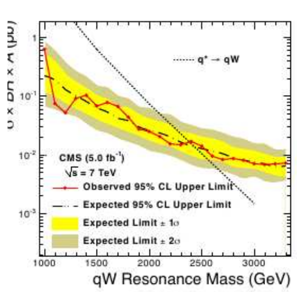


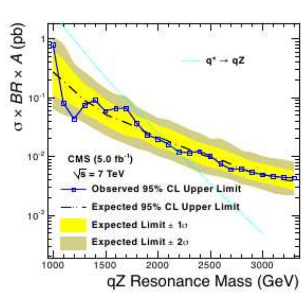


Searches with substructure tools

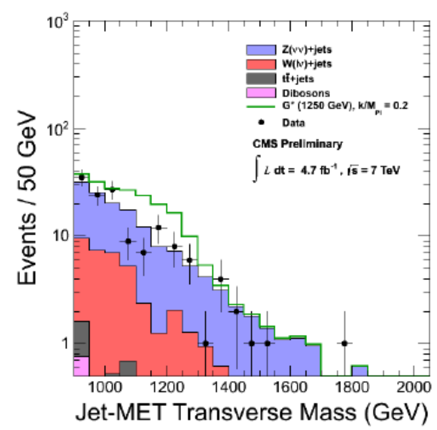
X → qV

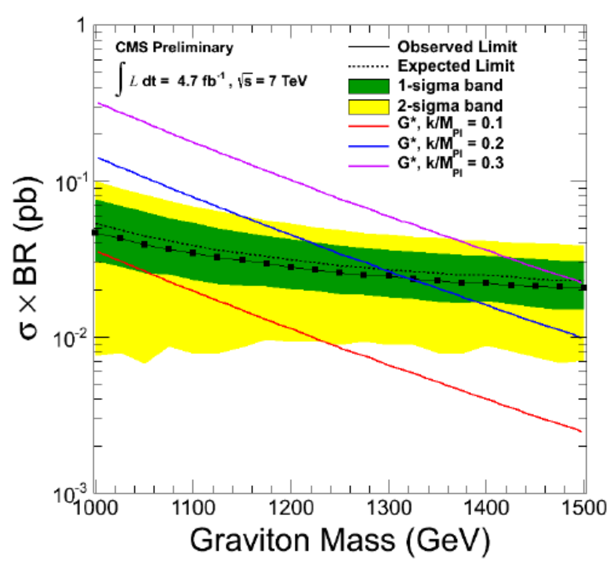


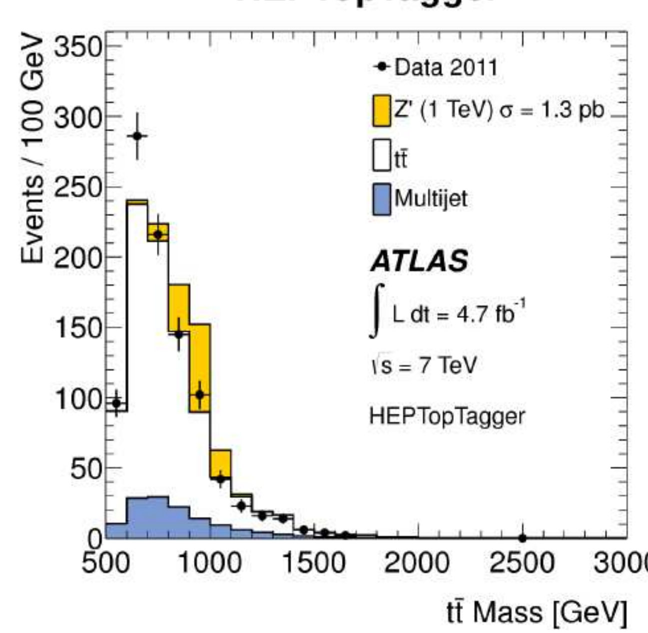


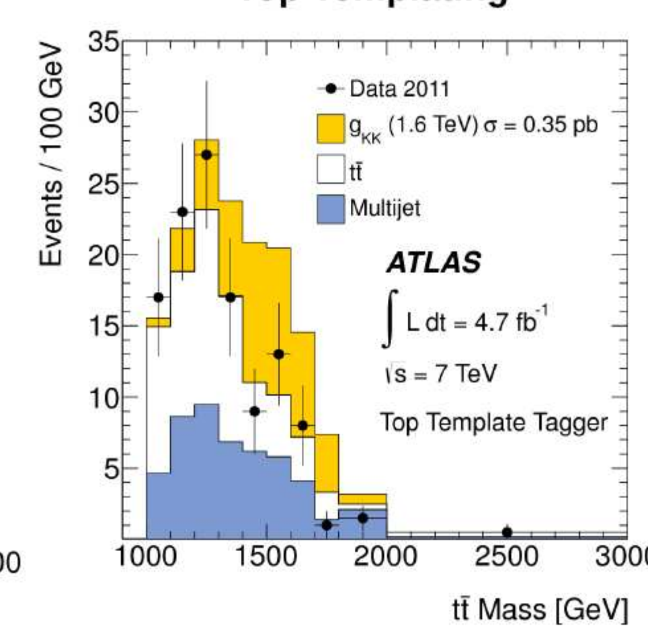


X → ZZ → 2ν2q

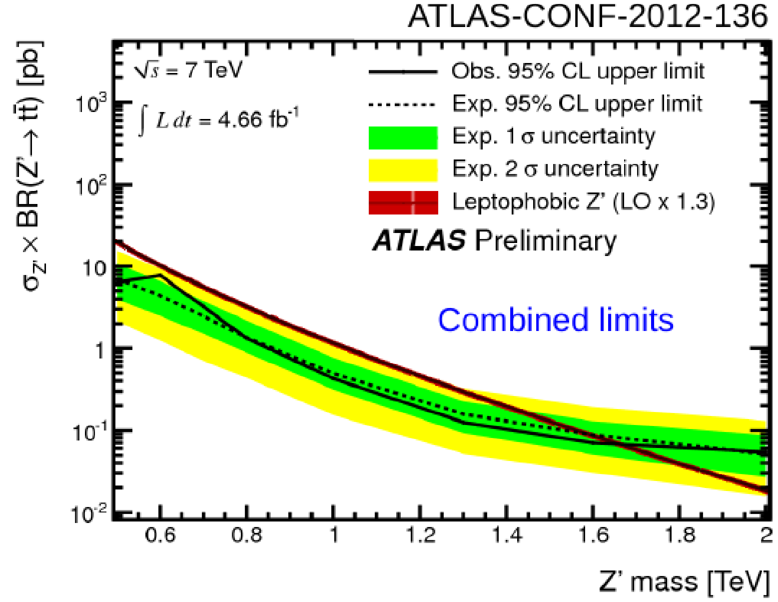








ATLAS-CONF-2012-136



A range of techniques being used for varied BSM scenarios

[Analytic] understanding

Work in progress with Dasgupta
Fregoso & Marzani

What do we know currently?

Boost 2010 proceedings:

The [Monte Carlo] findings discussed above indicate that while [pruning, trimming and filtering] have qualitatively similar effects, there are important differences. For our choice of parameters, pruning acts most aggressively on the signal and background followed by trimming and filtering.

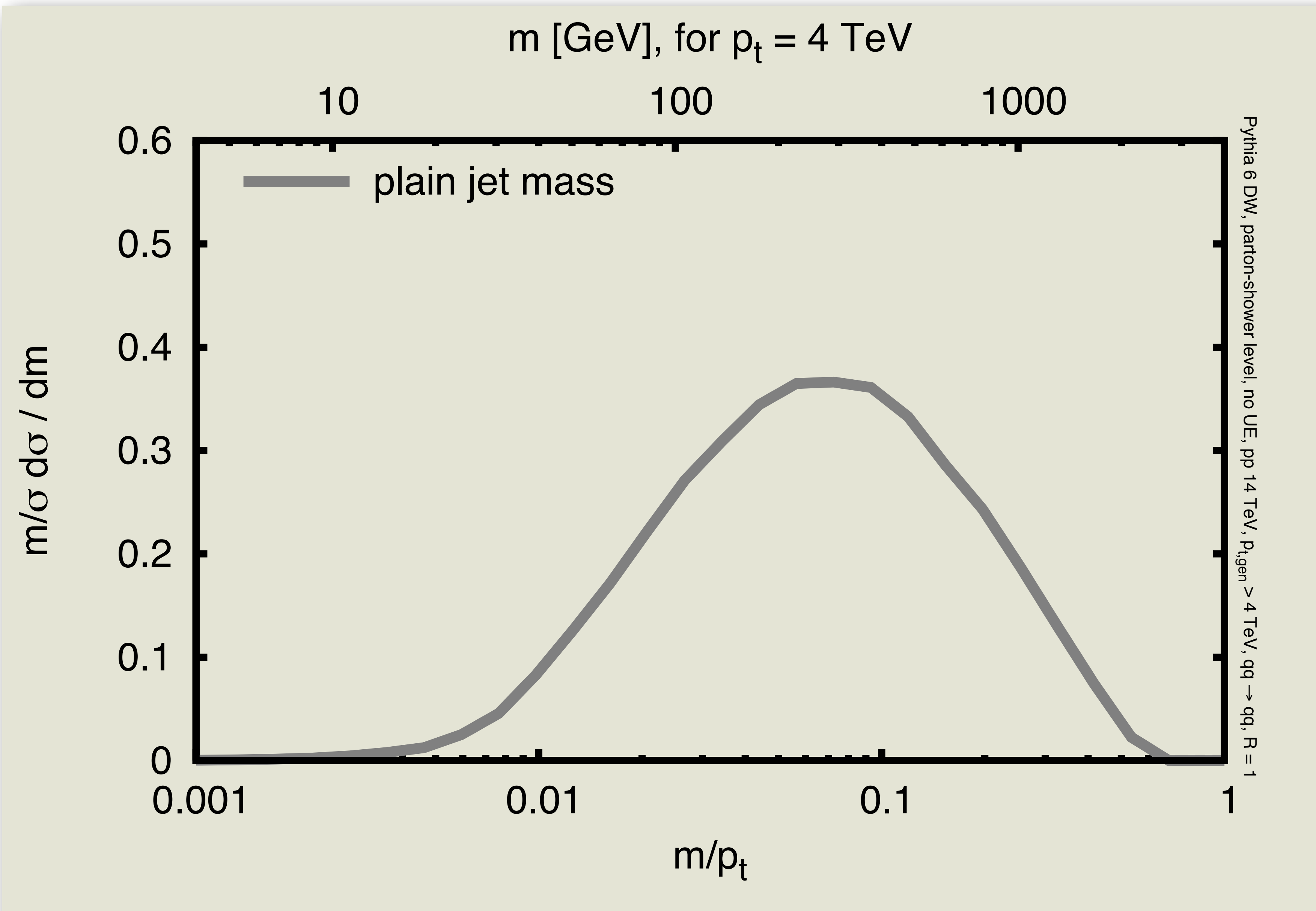
At the time:

- No clear picture of why the taggers might be similar or different
- No clear picture of how the parameter choices affect the taggers

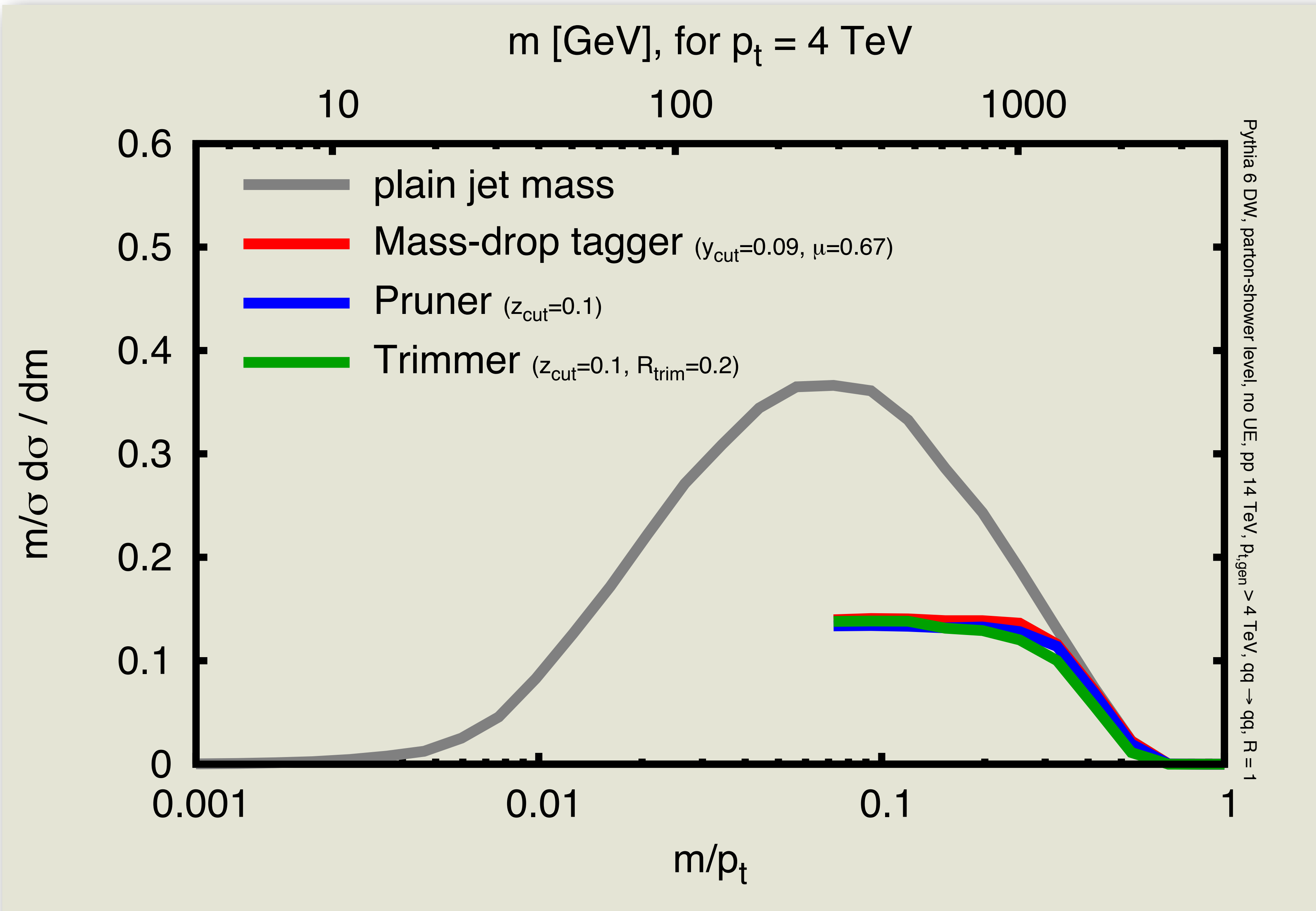
Today:

- I'll show a selection of **preliminary** lessons from studies for background jets in progress with Dasgupta, Fregoso and Marzani

The “right” MC study can already be instructive
(testing on background [quark] jets)

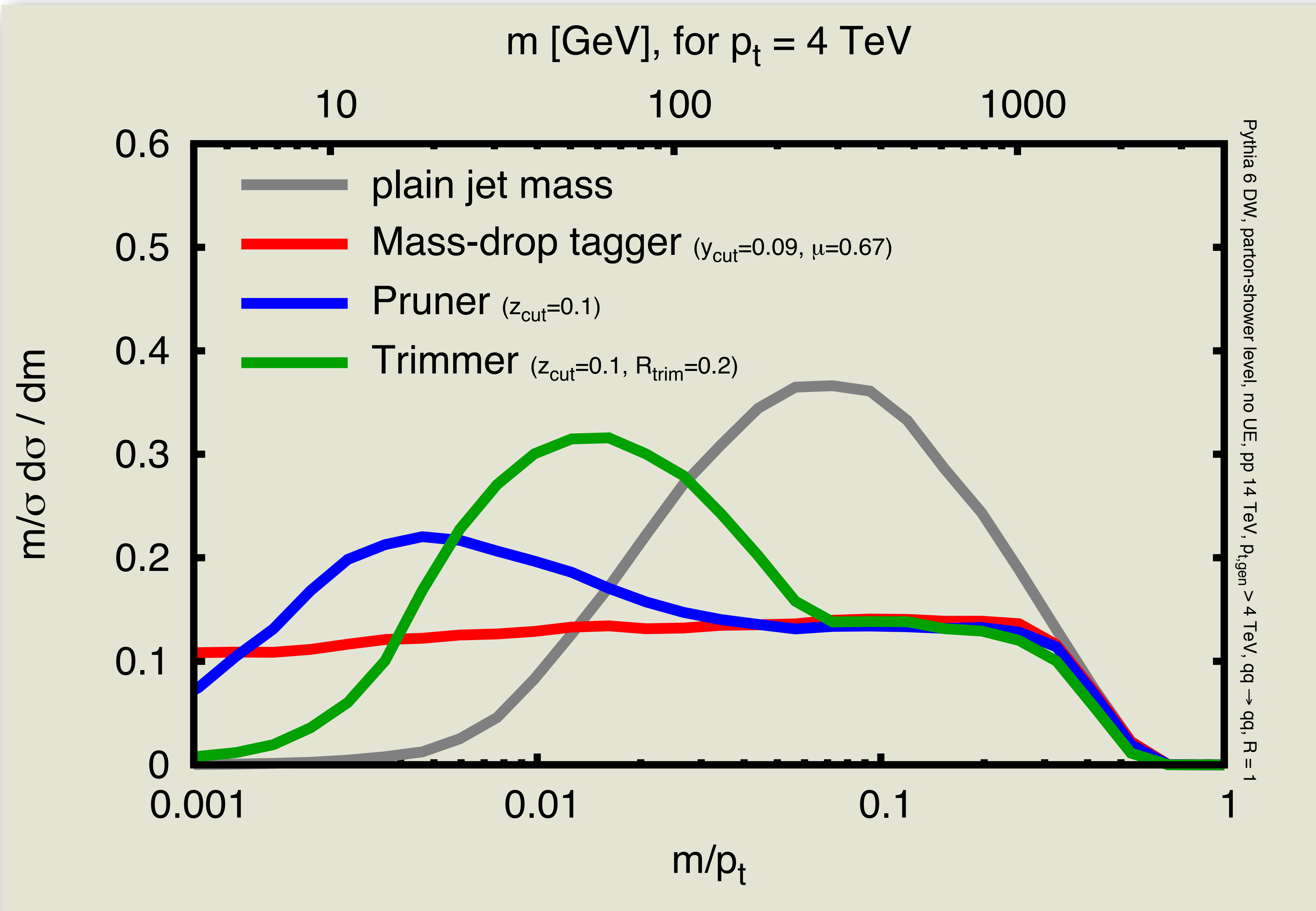


The "right" MC study can already be instructive
(testing on background [quark] jets)



Different taggers are apparently quite similar

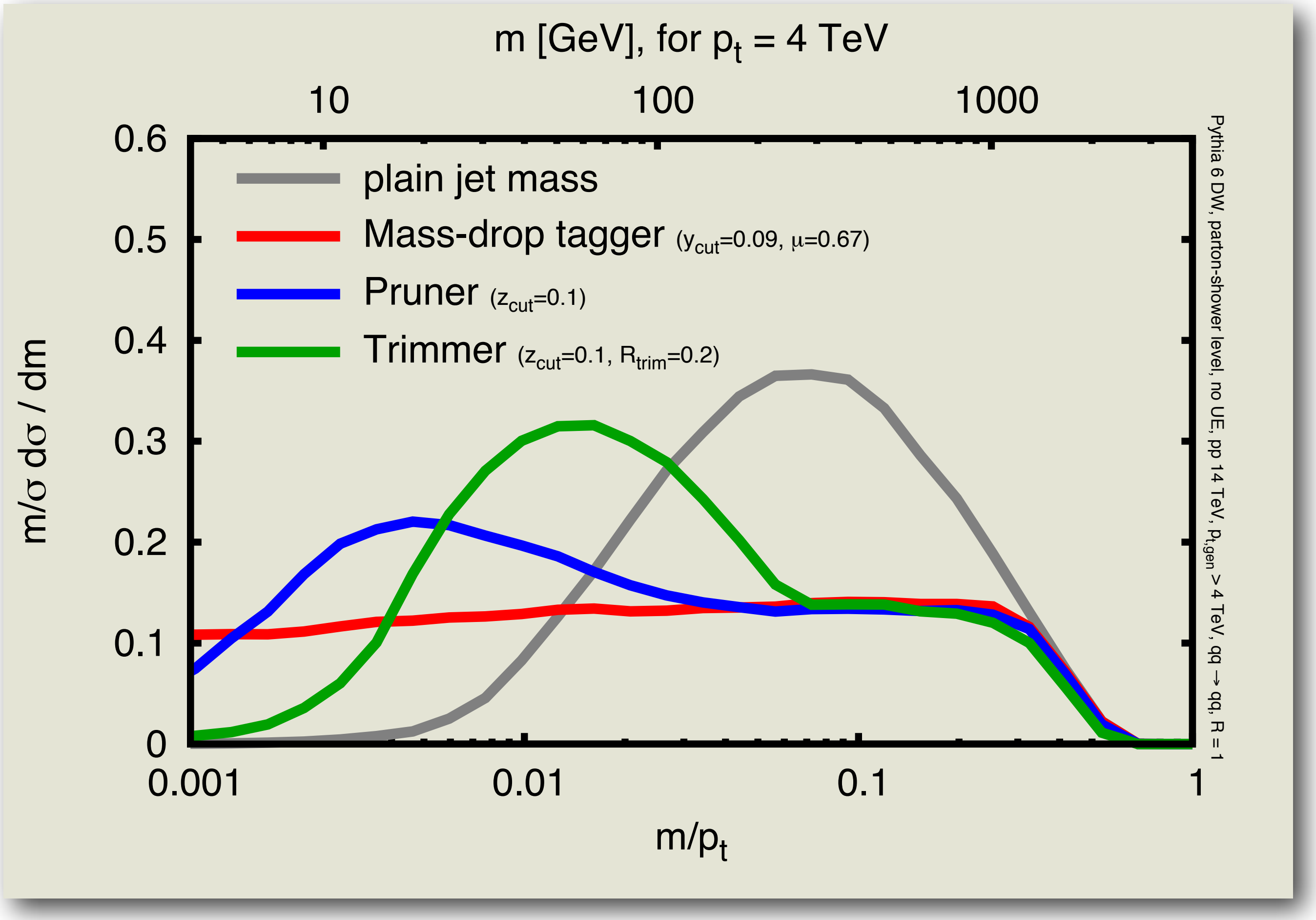
The "right" MC study can already be instructive
(testing on background [quark] jets)



But only for a limited range of masses

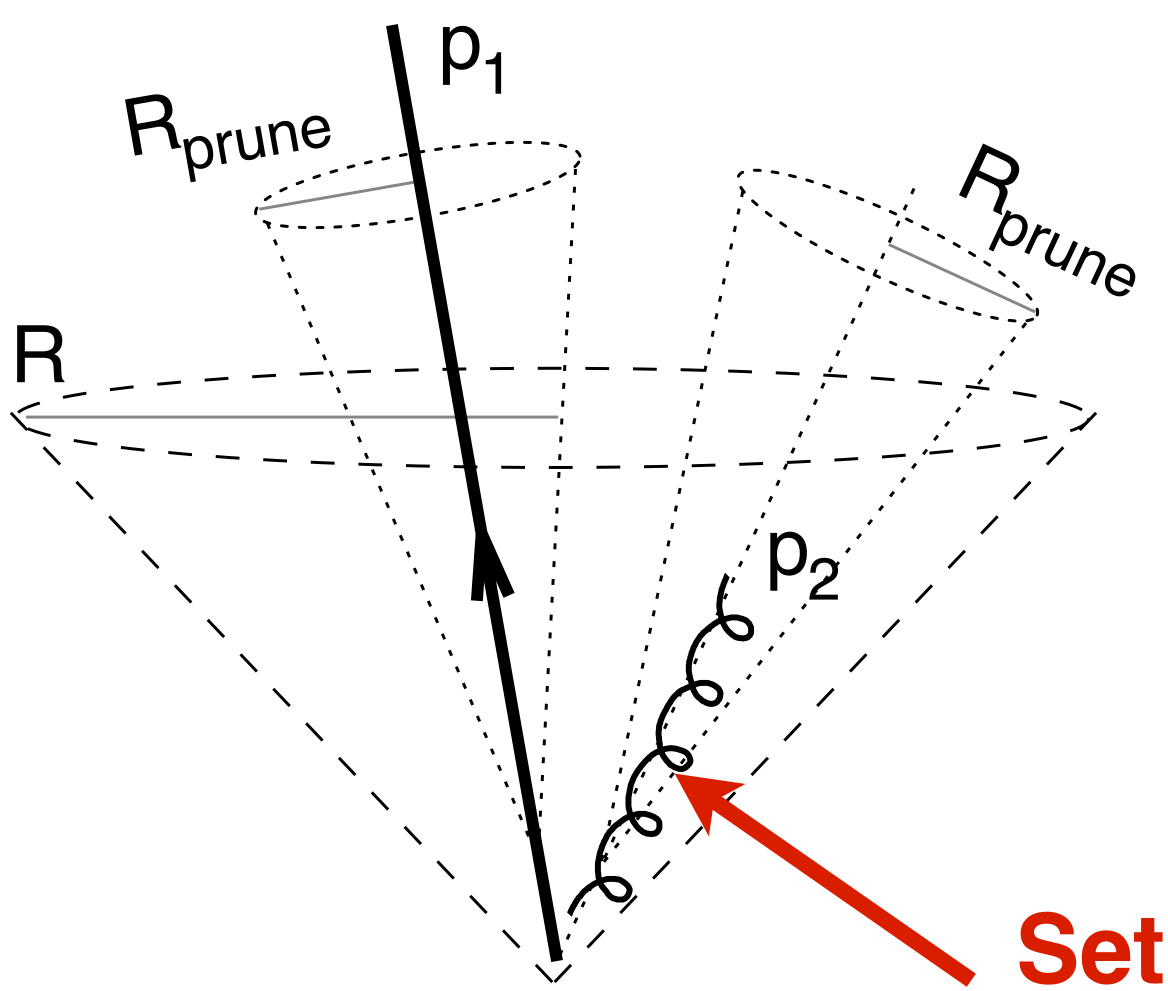
Do we care about such differences?

- Think data-driven backgrounds: kinks can seriously mess with you [especially if you got used to their being absent, e.g. from moderate p_t tests]
- How do these structures depend on the z_{cut}, y_{cut}, R_{trim}, etc. parameters?
- Are these structures telling us something we might want to know about the taggers? E.g. how to improve them?



This calls for analysis and calculation

Pruning [7, 8] takes an initial jet, and from its mass deduces a pruning radius $R_{\text{prune}} = R_{\text{fact}} \cdot \frac{2m}{p_t}$, where R_{fact} is a parameter of the tagger. It then reclusters the jet and for every clustering step, involving objects a and b, it checks whether $\Delta_{ab} > R_{\text{prune}}$ and $\min(p_{ta}, p_{tb}) < z_{\text{cut}} p_{t,(a+b)}$, where z_{cut} is a second parameter of the tagger. If so, then the softer of the a and b is discarded. Otherwise a and b are recombined as usual. Clustering then proceeds with the remaining objects, applying the pruning check at each stage.



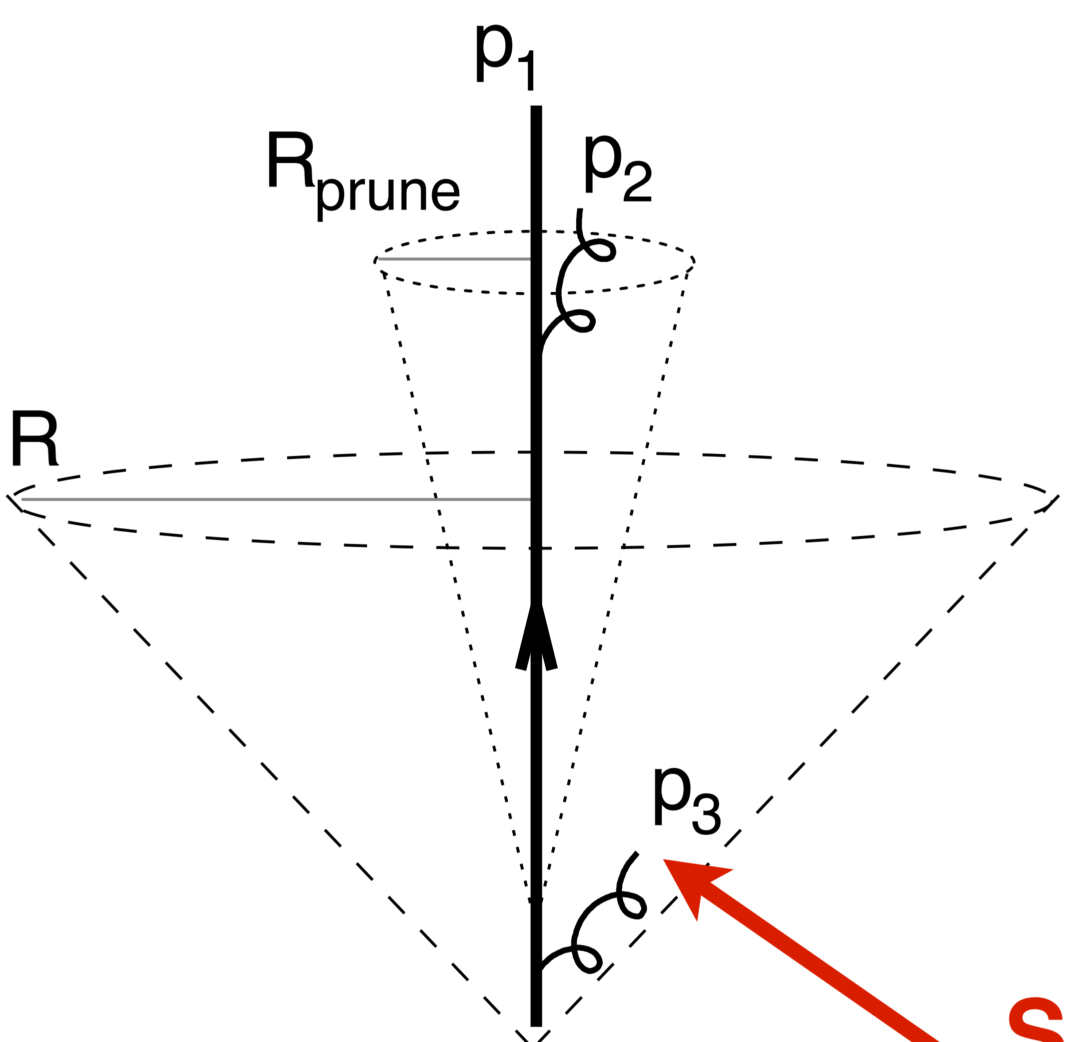
What pruning is meant to do:

Choose an R_{prune} such that different hard prongs (p_1, p_2) end up in different hard subjets.

Discard any softer radiation.

Sets pruning radius, & hard enough to end up as subjet

Pruning [7, 8] takes an initial jet, and from its mass deduces a pruning radius $R_{\text{prune}} = R_{\text{fact}} \cdot \frac{2m}{p_t}$, where R_{fact} is a parameter of the tagger. It then reclusters the jet and for every clustering step, involving objects a and b, it checks whether $\Delta_{ab} > R_{\text{prune}}$ and $\min(p_{ta}, p_{tb}) < z_{\text{cut}} p_{t,(a+b)}$, where z_{cut} is a second parameter of the tagger. If so, then the softer of the a and b is discarded. Otherwise a and b are recombined as usual. Clustering then proceeds with the remaining objects, applying the pruning check at each stage.



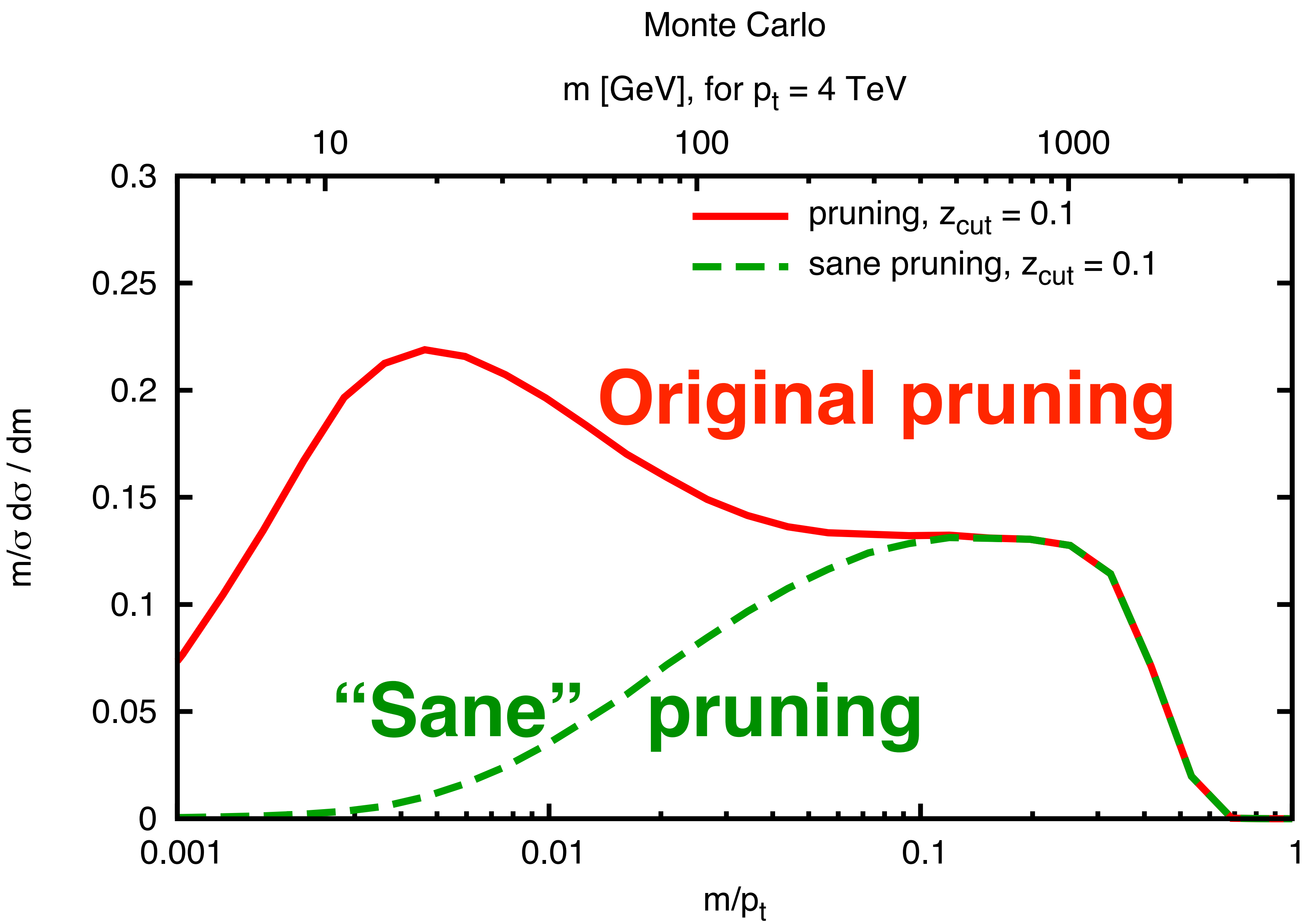
What pruning sometimes does

Chooses R_{prune} based on a soft p_3 (dominates total jet mass), and leads to a single narrow subjet whose mass is also dominated by a soft emission (p_2, within R_{prune} of p_1, so not pruned away).

Sets pruning radius, but gets pruned away

A simple fix: “sane” pruning

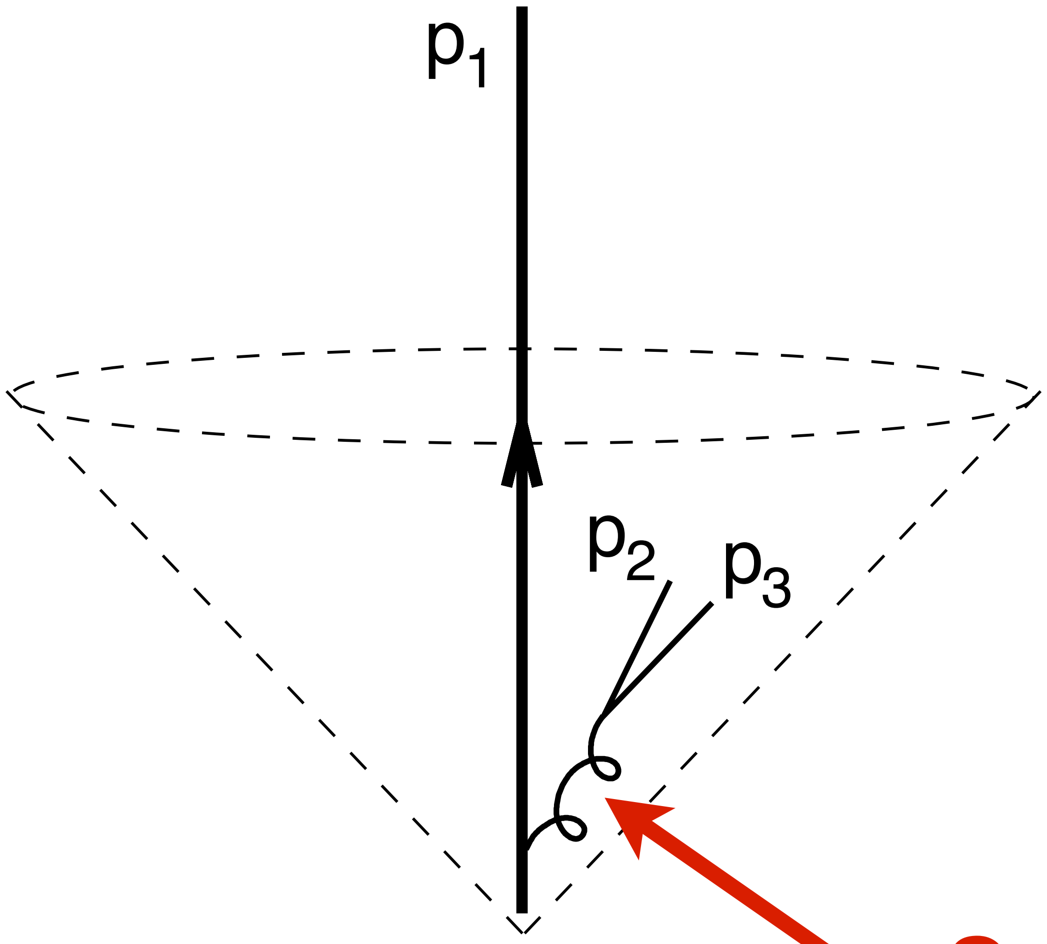
Require at least one successful merging with $\Delta R > R_{prune}$ and $z > z_{cut}$



“sane” pruning is effectively placing an isolation cut on radiation around the tagged object

Mass drop tagger

1. Break the jet j into two subjets by undoing its last stage of clustering. Label the two subjets j_1, j_2 such that $m_{j_1} > m_{j_2}$.
2. If there was a significant mass drop, $m_{j_1} < \mu m_j$, and the splitting is not too asymmetric, $y = \min(p_{tj_1}^2, p_{tj_2}^2)\Delta R_{j_1 j_2}^2 / m_j^2 > y_{\text{cut}}$, then deem j to be the tagged jet.
3. Otherwise redefine j to be equal to j_1 and go back to step 1 (unless j consists of just a single particle, in which case the original jet is deemed untagged).



<u>What MDT does wrong:</u>

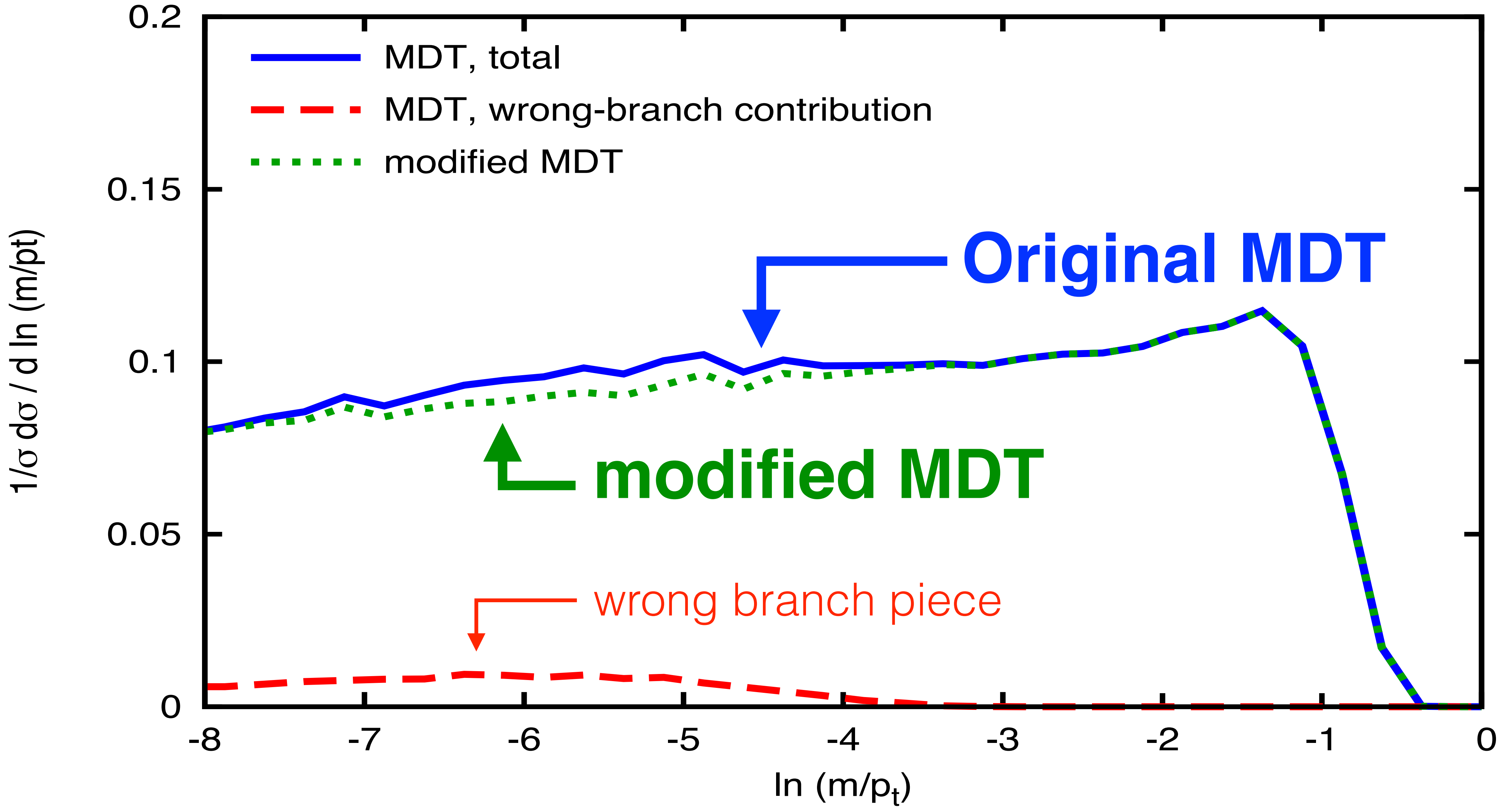
Follows a soft branch ($p_2+p_3 < y_{cut}\ p_{jet}$) with "accidental" small mass, when the "right" answer was that the (massless) hard branch had no substructure

Subjet is soft, but has more substructure than hard subjet

A simple fix for “modified” Mass Drop Tagger:

When recursing, follow branch with larger ($m^2+p_t^2$)

(rather than the one with larger m)



What about analytic calculations of the taggers?

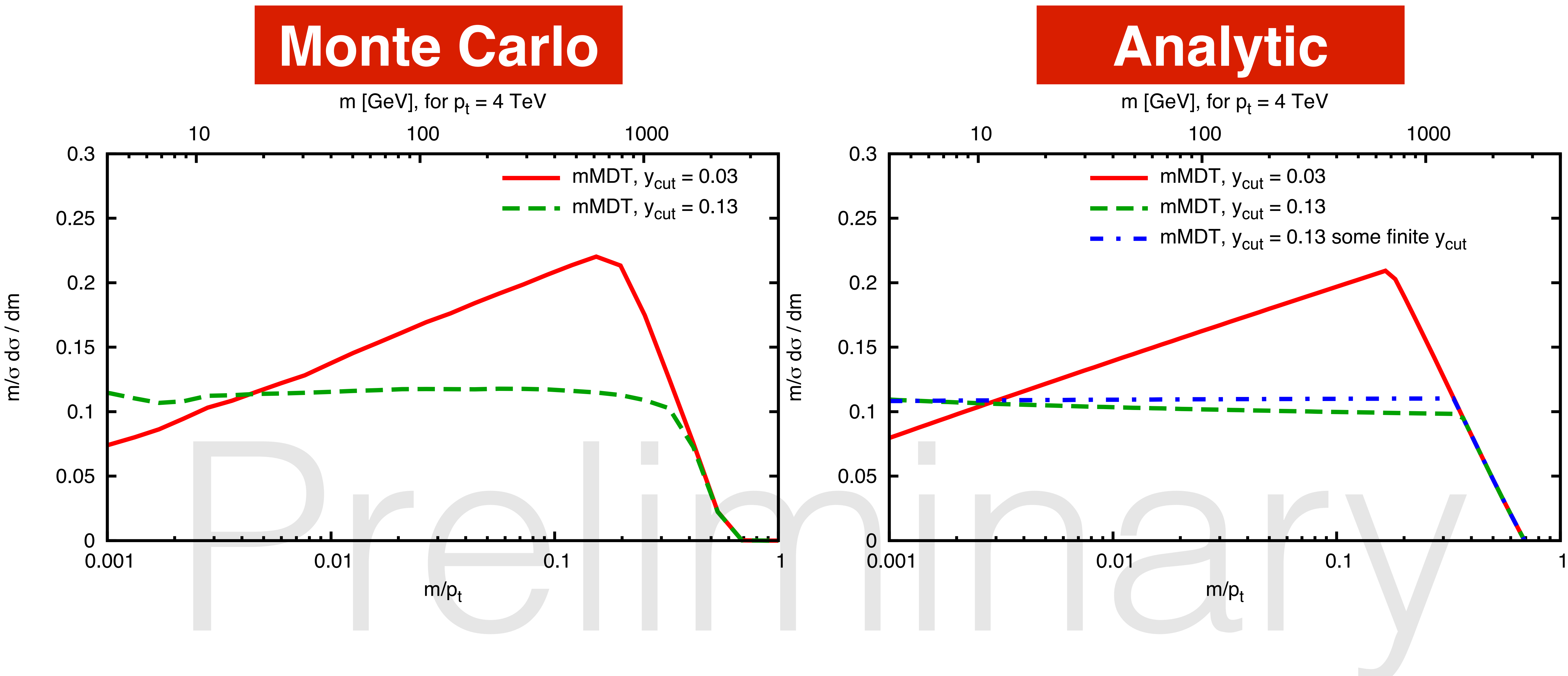
Simpler　　　　　　　　　　　　　　　　　More complex

Dasgupta et al (full NLL, including non-global part)
Chien et al (partial NNLL)
Jouttenus et al (they say full NNLL; I remain to be convinced!)

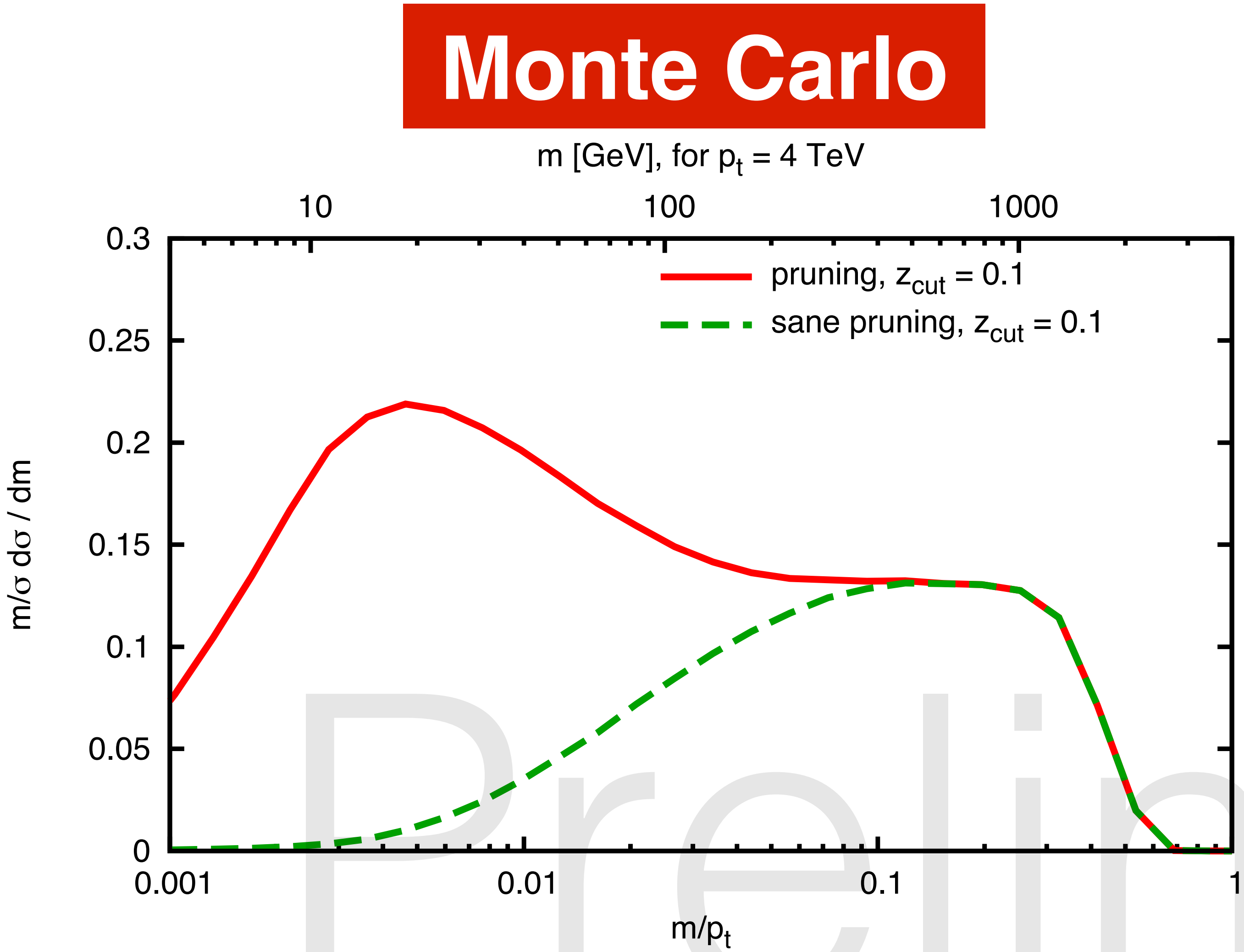
mMDT > plain mass ~ trimming > sane pruning > full pruning ≫ ~~MDT~~

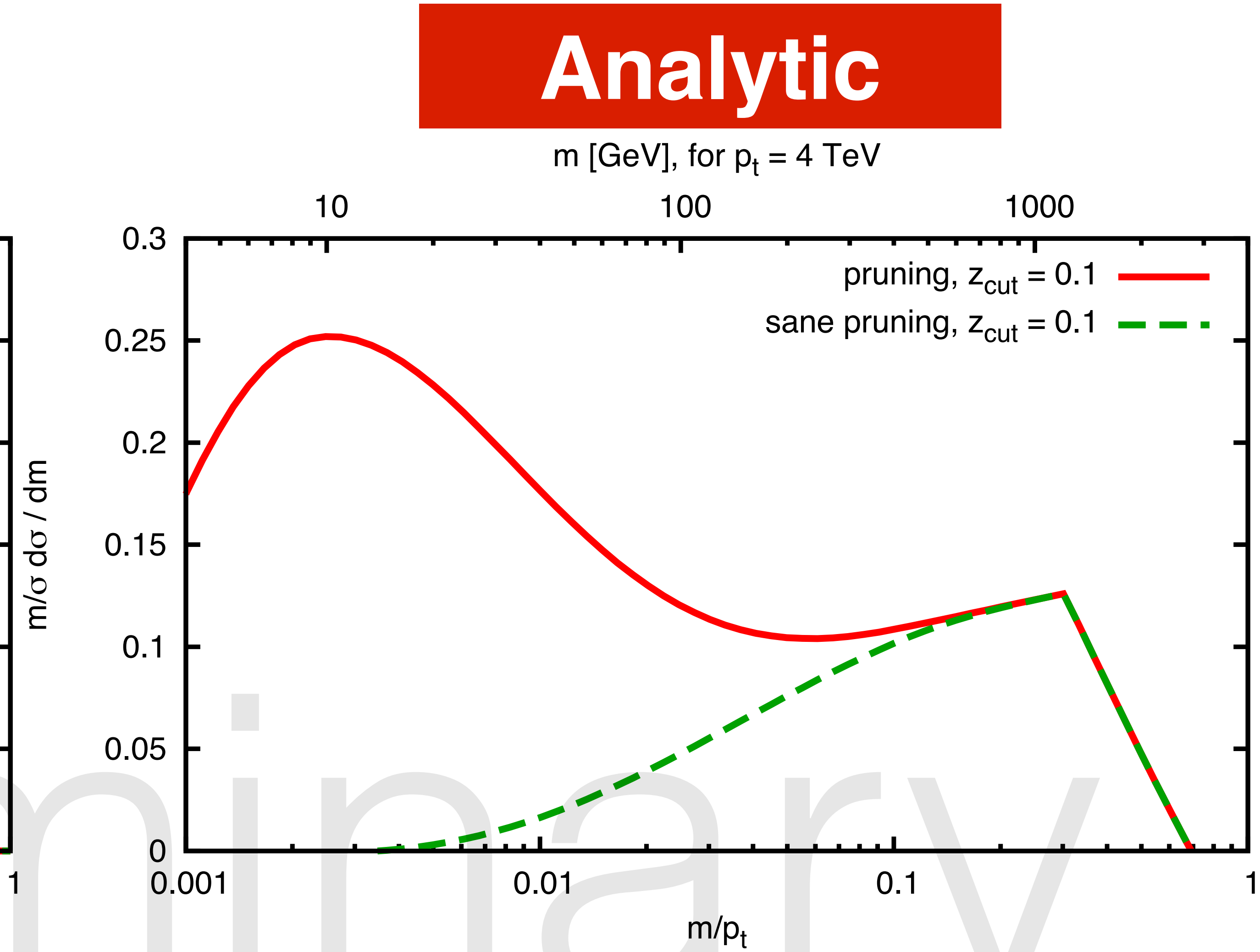
Dasgupta, Fregoso, Marzani & GPS, forthcoming:
LL in all cases, plus some subleading logs
[NB: LL doesn't mean the same thing in all cases!)

Modified Mass Drop Tagger

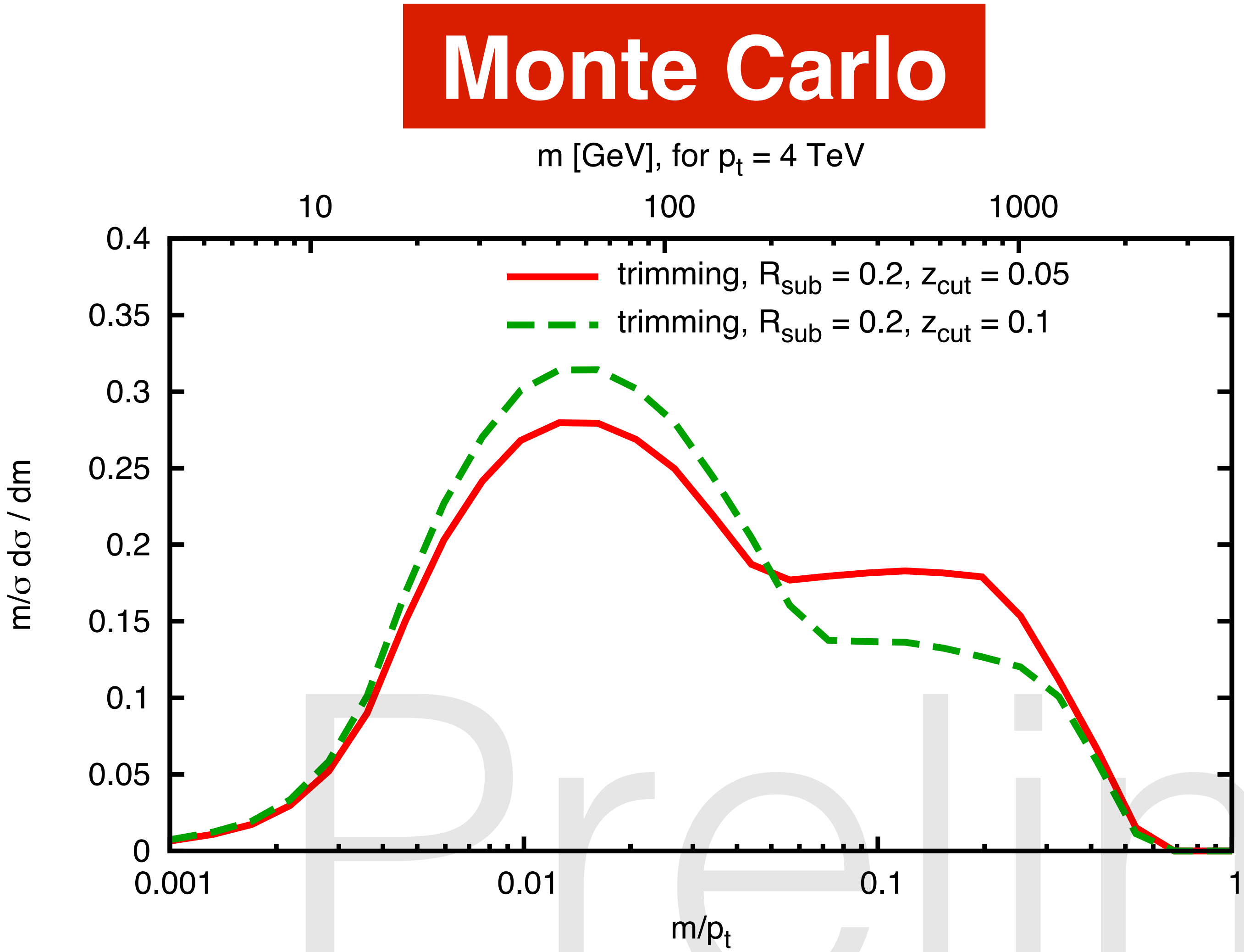


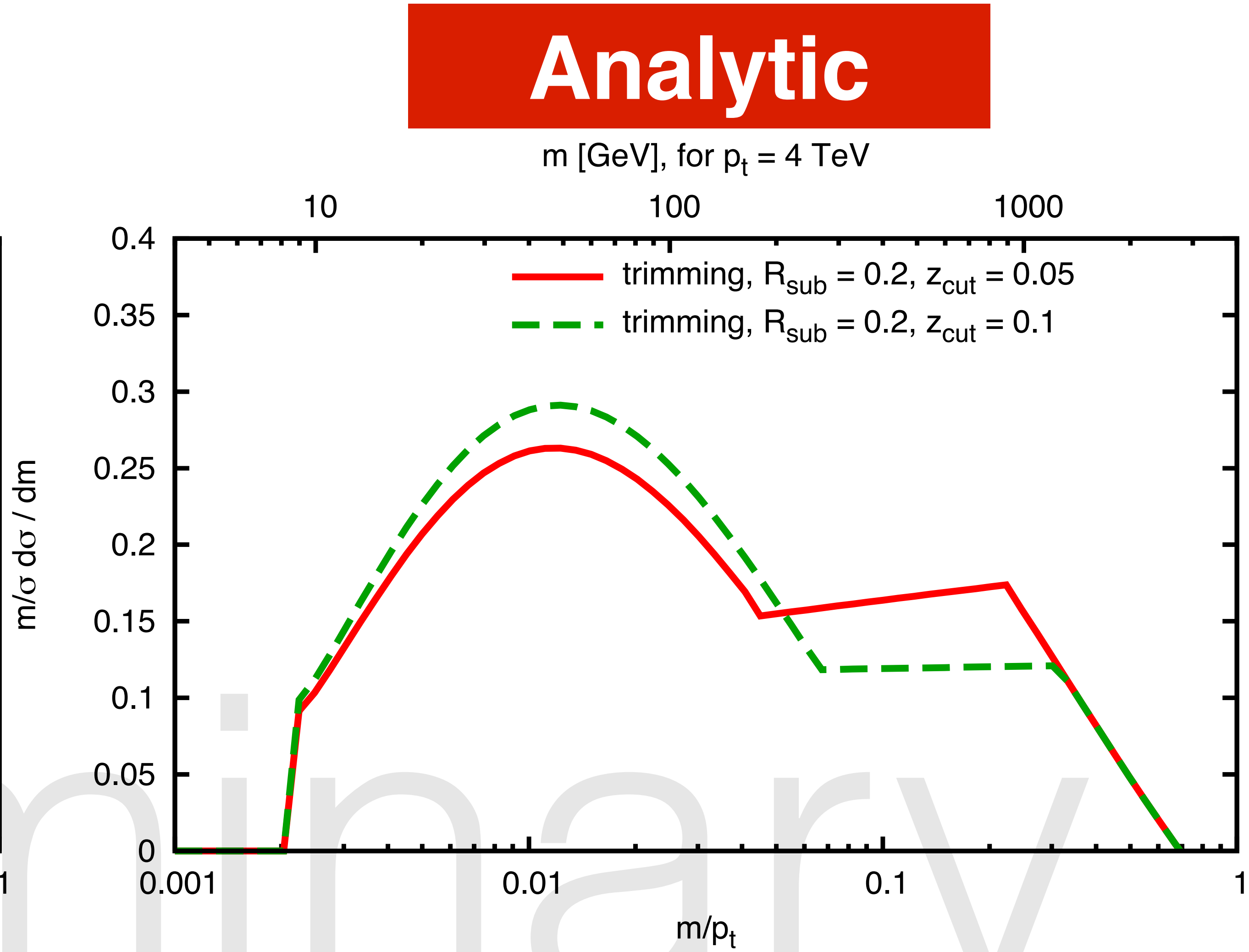
[mMDT is closest we have to a scale-invariant tagger, though exact behaviour depends on q/g fractions]





Preliminary


Monte Carlo
m [GeV], for p_t = 4 TeV
10
100
1000
trimming, R_{sub} = 0.2, z_{cut} = 0.05
trimming, R_{sub} = 0.2, z_{cut} = 0.1
m/σ dσ / dm
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.001
0.01
0.1
1
m/p_t


Analytic
m [GeV], for p_t = 4 TeV
10
100
1000
trimming, R_{sub} = 0.2, z_{cut} = 0.05
trimming, R_{sub} = 0.2, z_{cut} = 0.1
m/σ dσ / dm
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.001
0.01
0.1
1
m/p_t

Bottom line on “understanding”

- Taggers may be quite simple to write, but potentially involved to understand.

- Contrast this with p_t cuts for standard jet analyses – (mostly) simple

- Still, many taggers/groomers are within calculational reach.

- New “modified” Mass Drop Tagger is especially simple

- New “sane” pruning is also interesting – further investigation warranted...

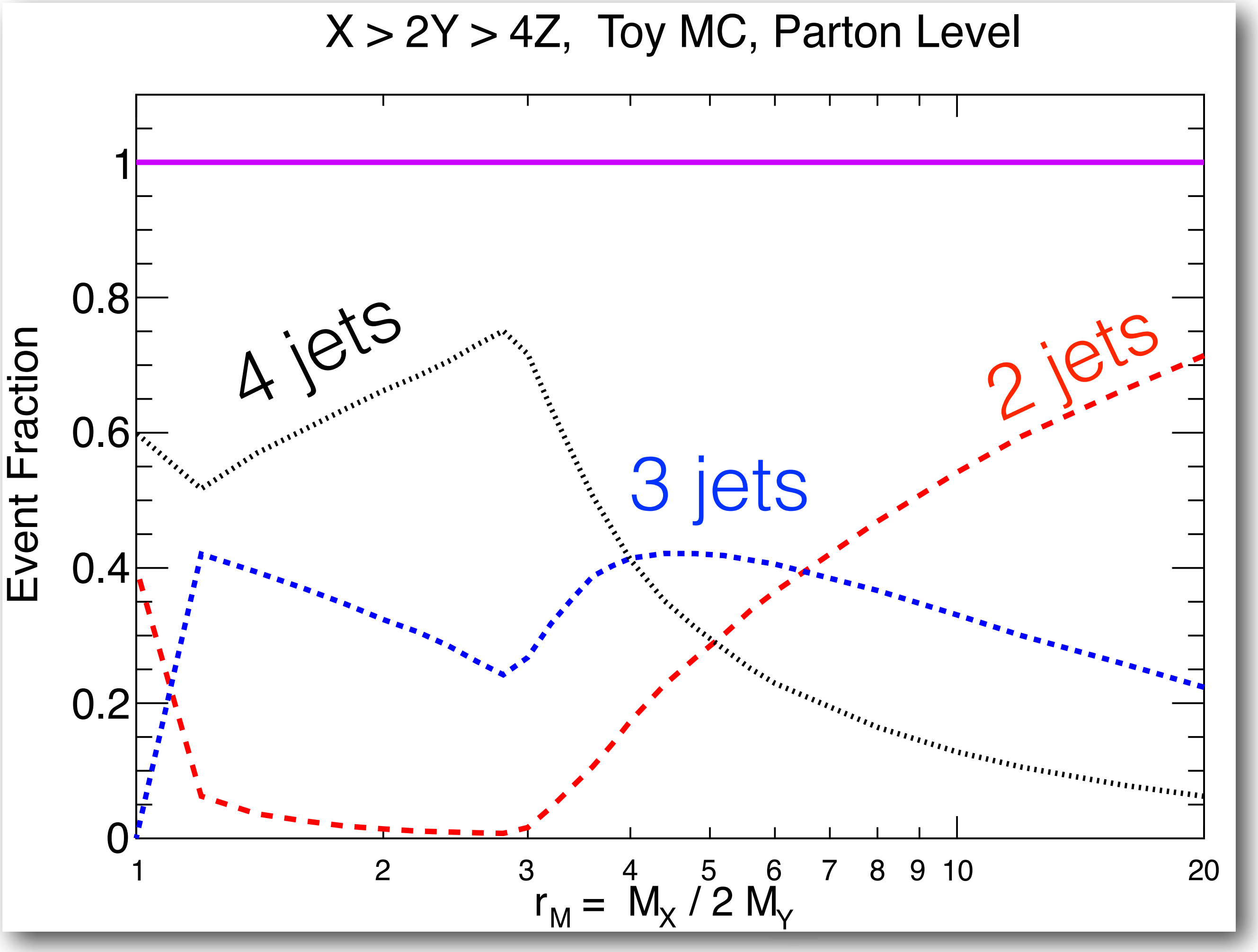
Scale invariant searches

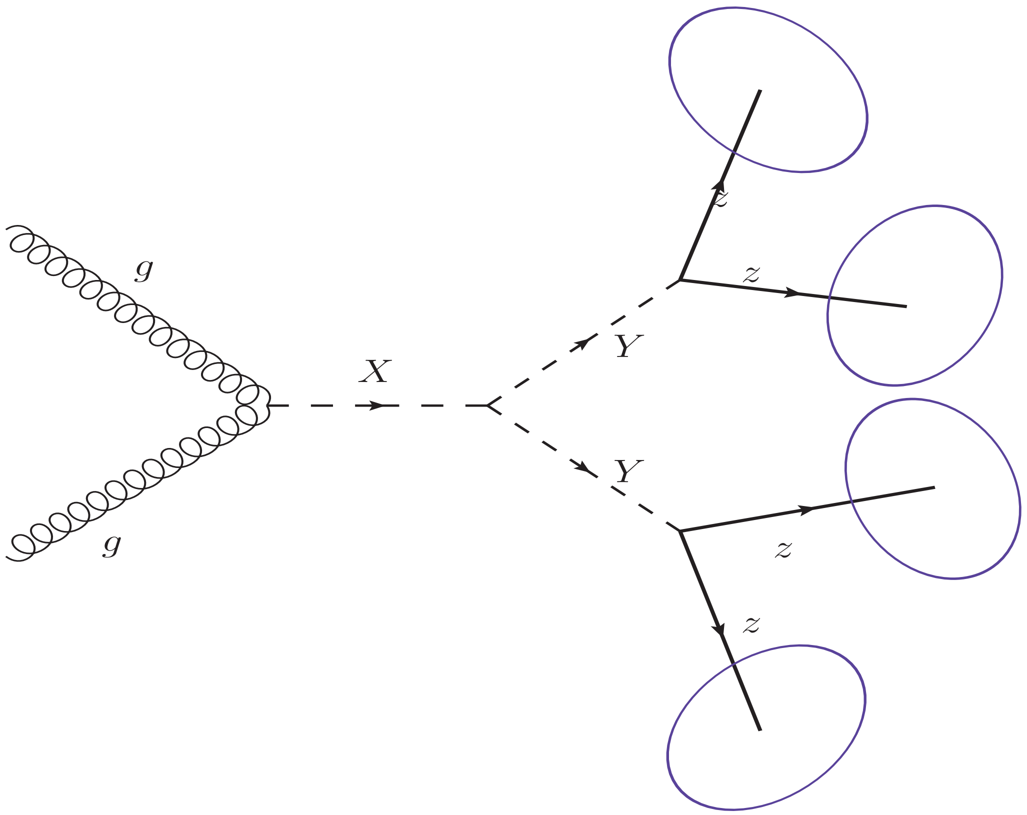
Work in progress with Gouzevitch, Oliveira, Rojo, Rosenfeld & Sanz

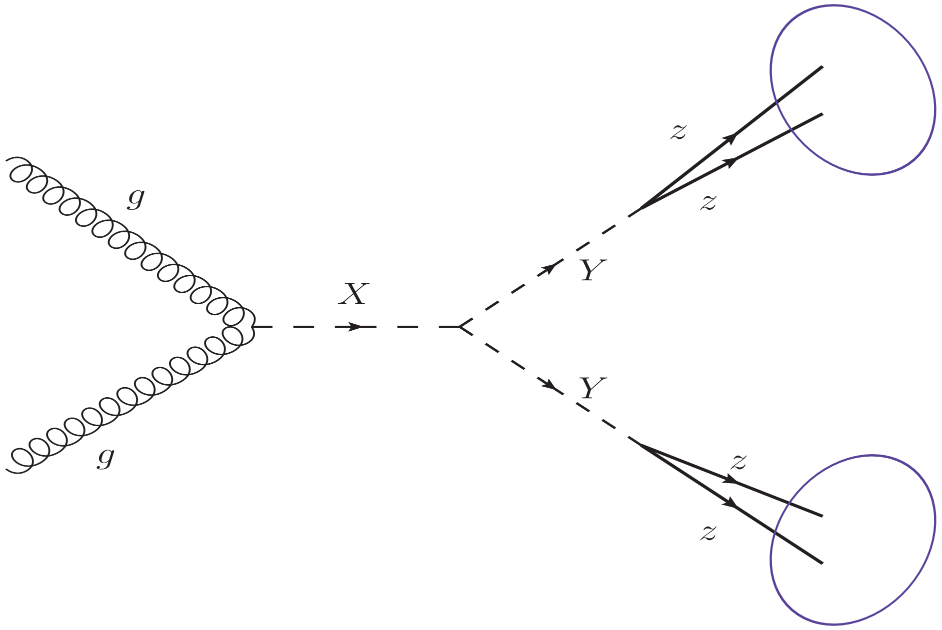
Experiments often have two distinct searches:

Resolved (small-R multi-jet)
Boosted (large-R fat-jet)

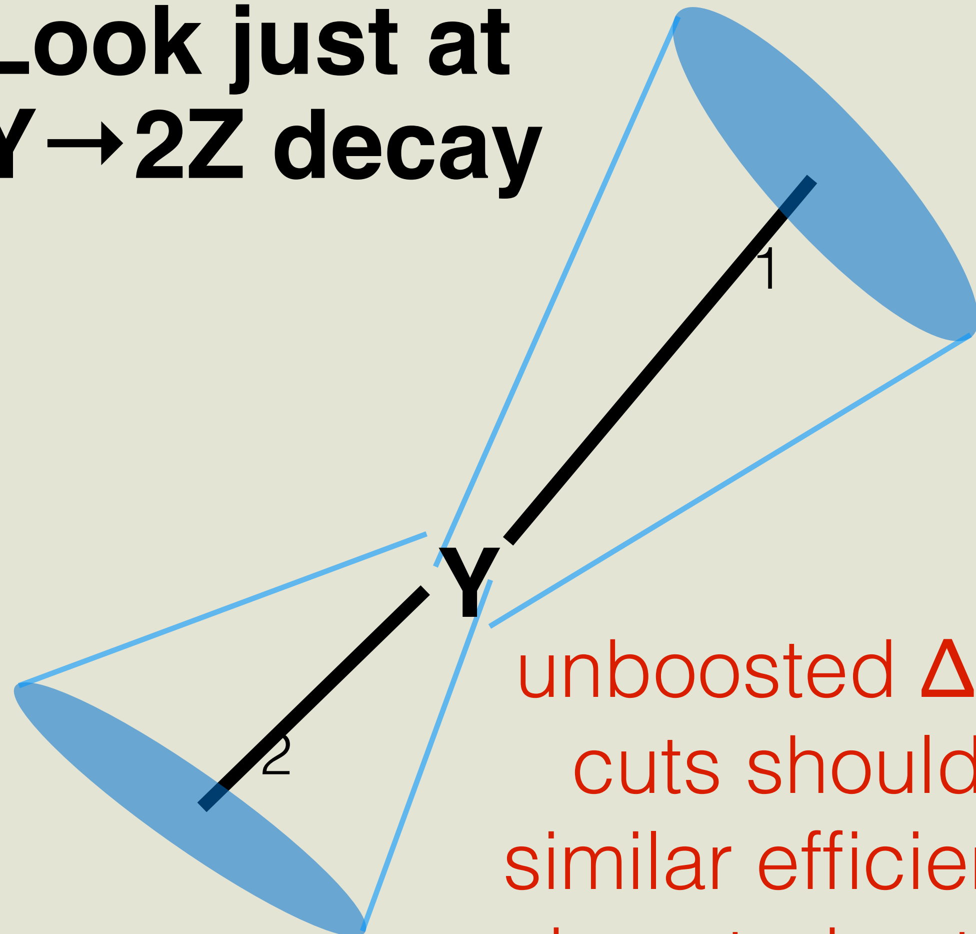
Can resolved and boosted analyses be consistently performed together?







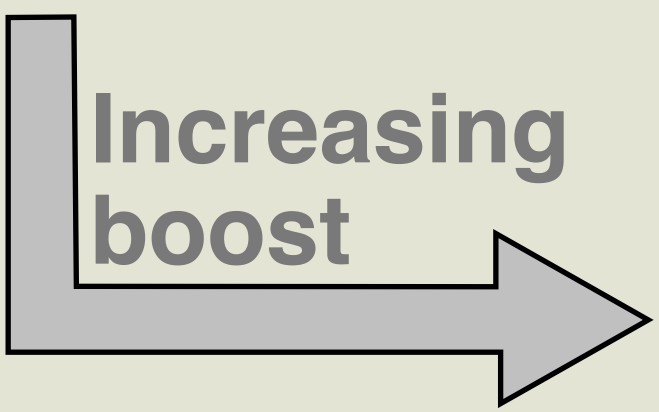
Look just at Y→2Z decay

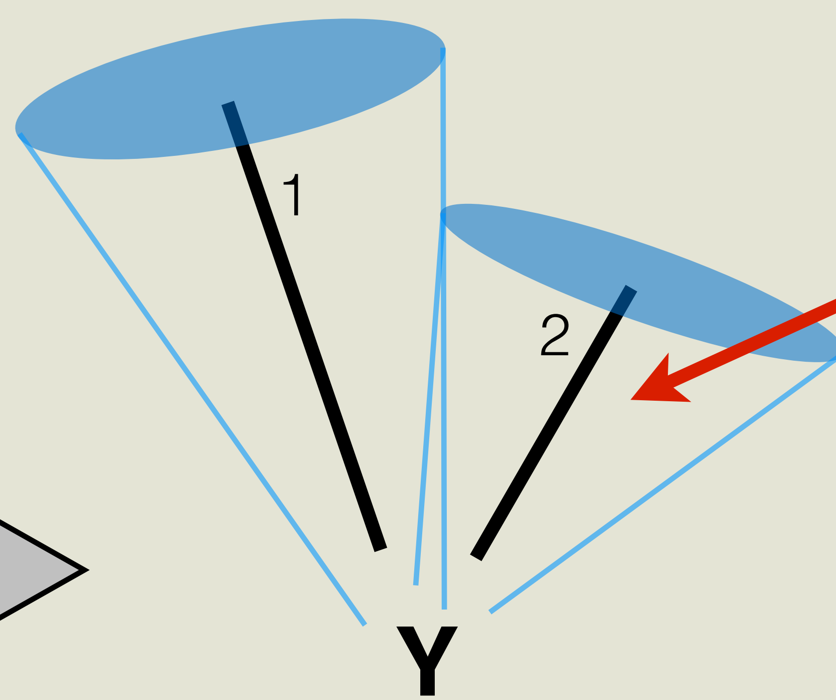


Key [simple] **idea:**

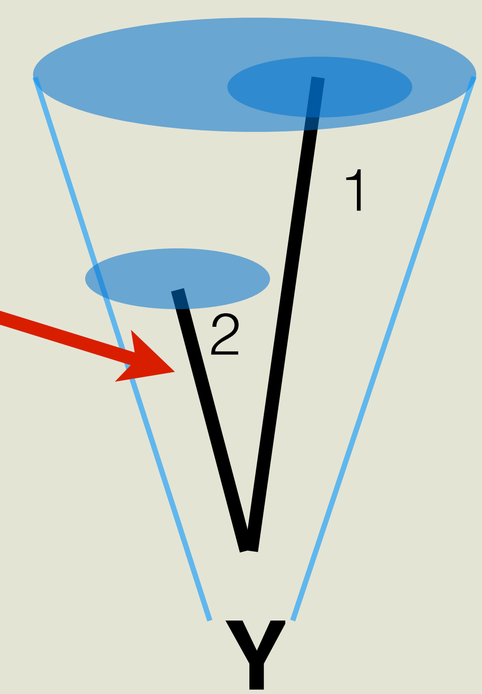
Cuts on resolved jets should mirror those on subjets inside fat jets

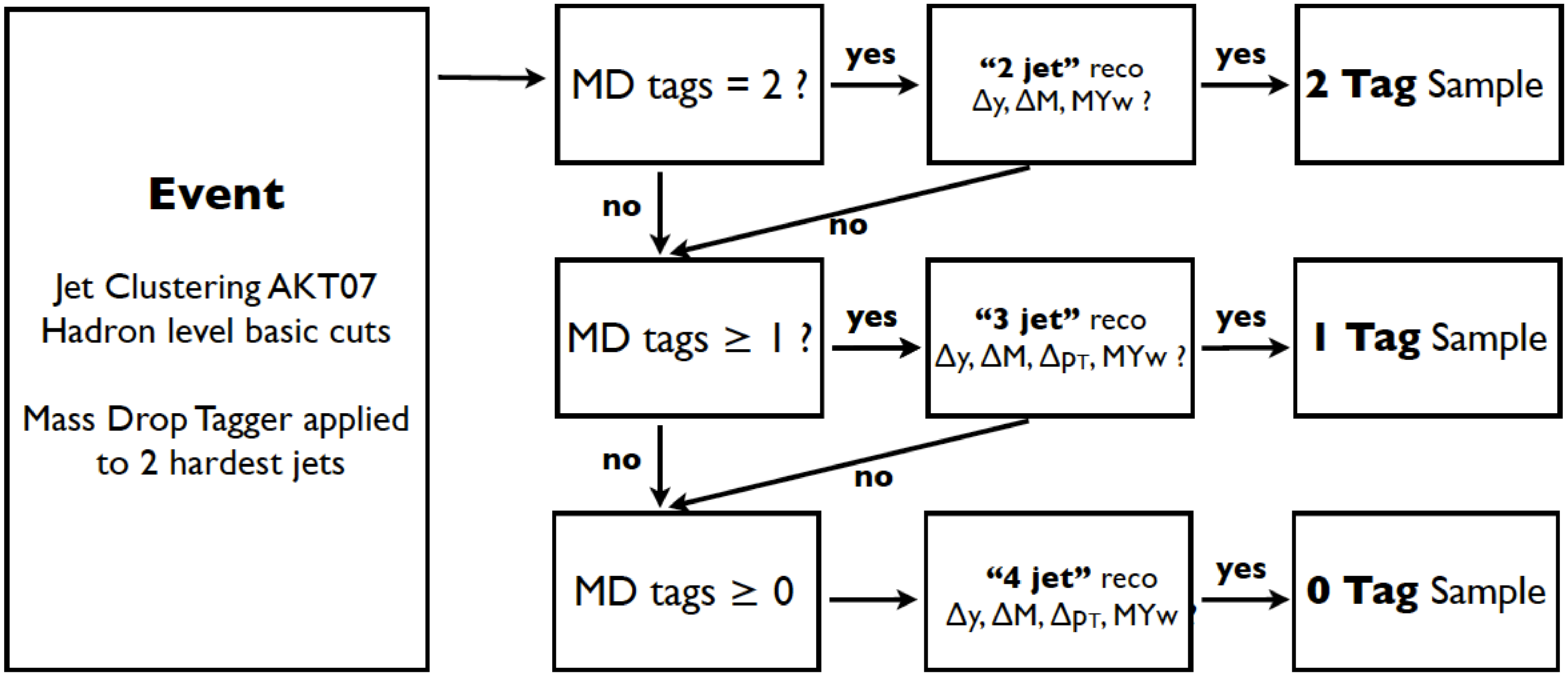
unboosted Δy_{12} & p_t cuts should have similar efficiencies to boosted cuts (y_{cut})





cuts on nearby jets should mirror subjets cuts, $p_{t2} \lesssim y_{cut}\, p_{t1}$





$$\left|\frac{2\,(m_{Y1}-m_{Y2})}{m_{Y1}+m_{Y2}}\right| \le f_m \qquad M_Y\,(1-f_m) \le m_{Y1}, m_{Y2} \le M_Y\,(1+f_m)$$

Traditional cuts on resolved jets

$$\Delta y \equiv |y_{Y1} - y_{Y2}| \le \Delta y_{\max} \qquad \Delta y \equiv |y_{Yi,1} - y_{Yi,2}| \le \Delta y^{\text{res}}_{\max}$$

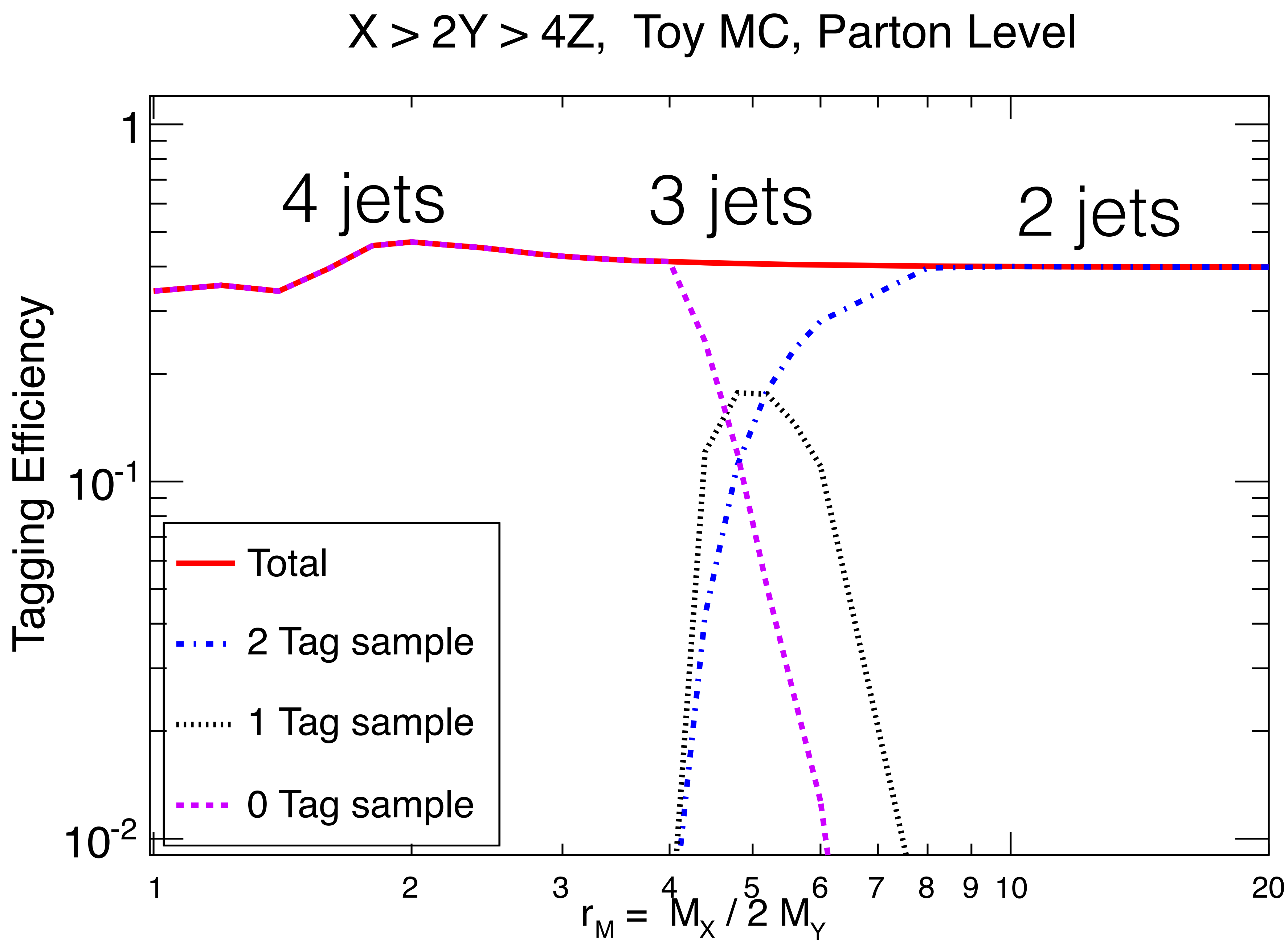
$$\Delta p_T \equiv p_T^{(1)} - p_T^{(2)} \ge (1-y_{\text{cut}})\,p_T^{(1)} \qquad \max\left(m_{Y,1}, m_{Y,2}\right) \le \mu \cdot m_Y$$

Boosted cuts on resolved jets

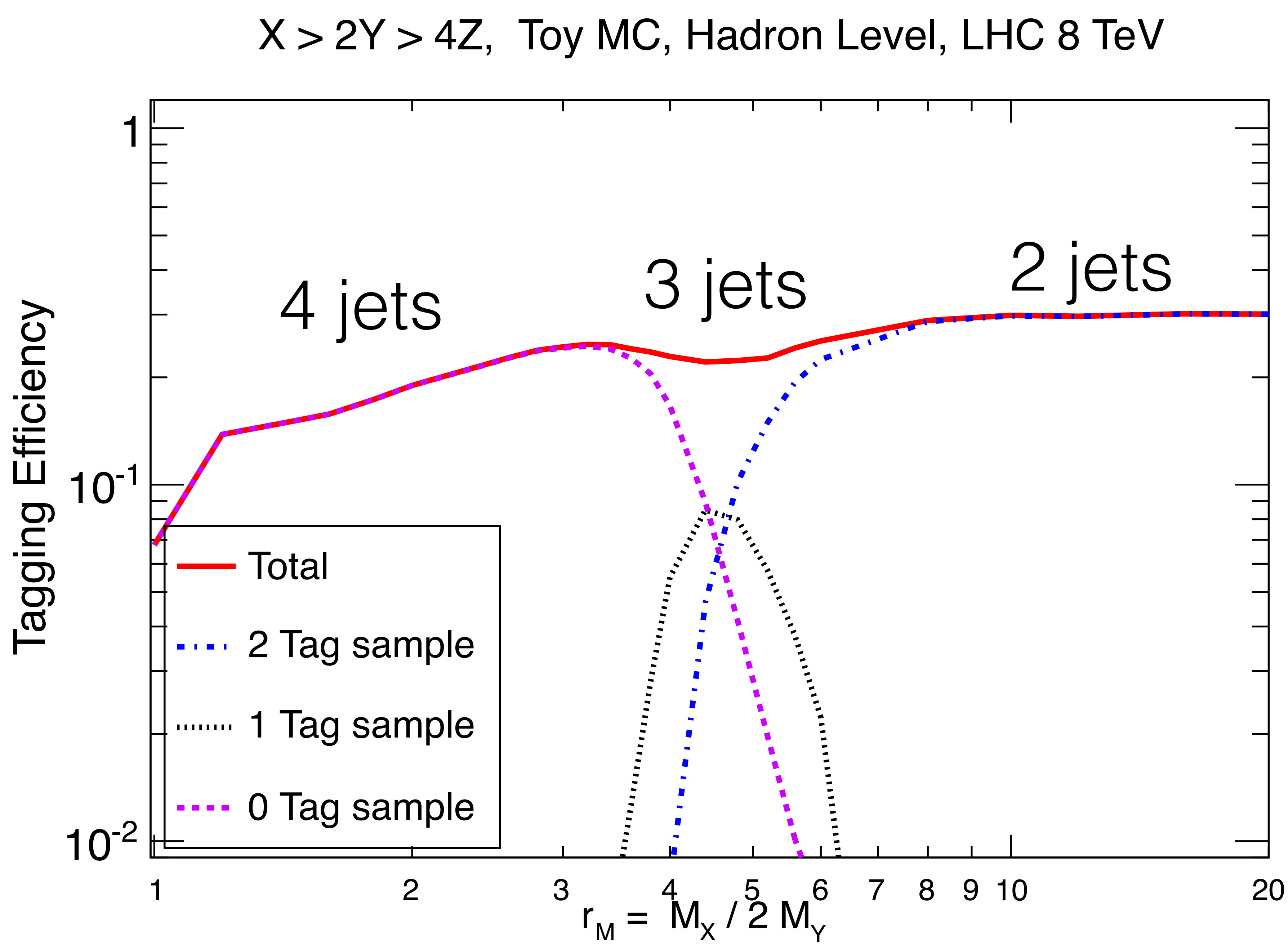
Jet Reconstruction					
R	R_{sj}	R_{f}	n_{filt}	μ	y_{cut}
0.5	1.3	0.3	3	0.67	0.09

Basic cuts		
$p_T^{\min}$	$\lvert y_{\max}\rvert$	$H_T^{\min}$
25 GeV	5.0	100 GeV

Quality requirements			
M_Y	$\Delta y_{\max}$	$\Delta y^{\text{res}}_{\max}$	f_m
125 GeV	1.3	1.5	0.15

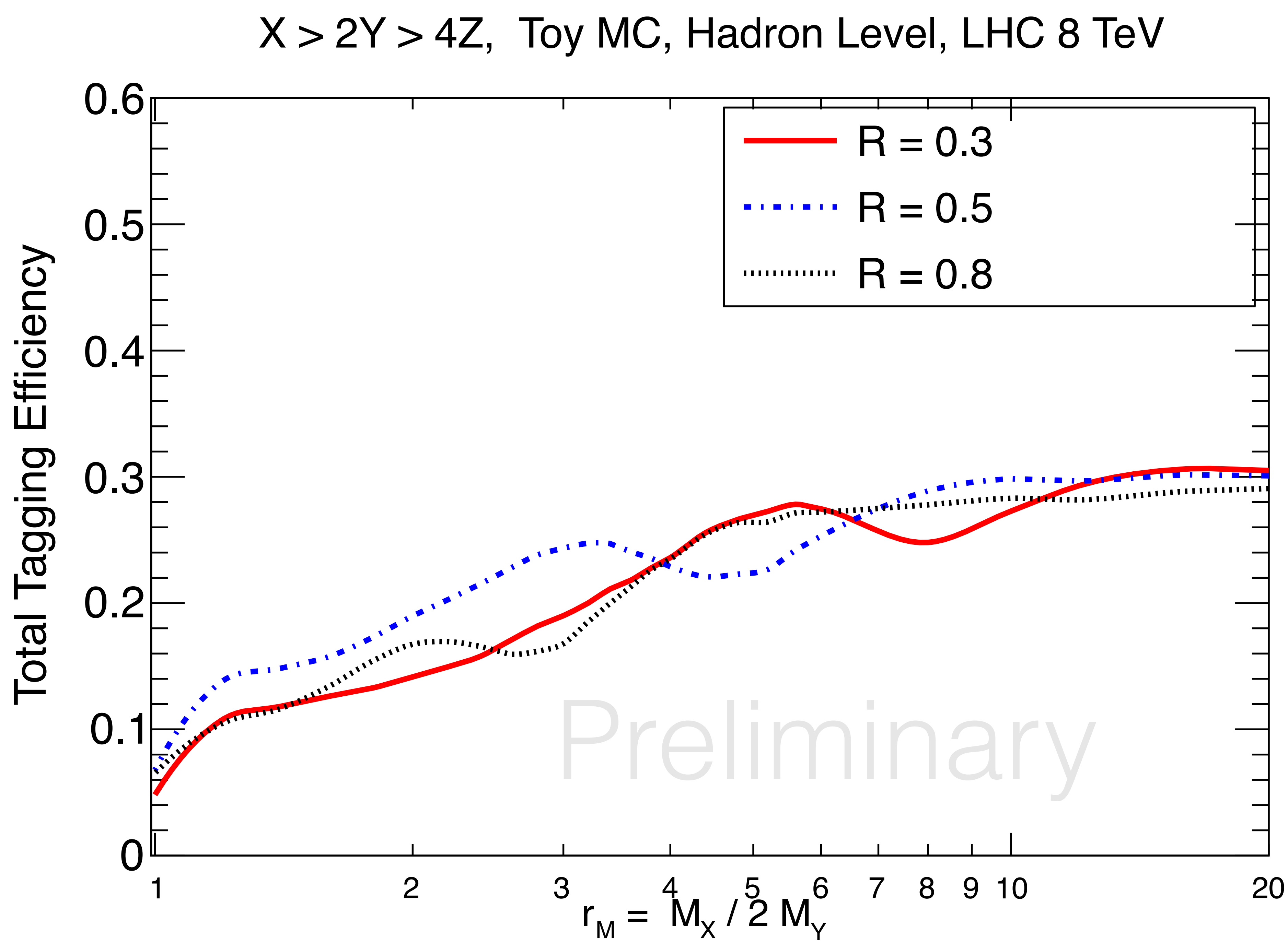


Efficiency roughly stable as X→2Y→4Z goes from 4→3→2 jets



Efficiency roughly stable as X→2Y→4Z goes from 4→3→2 jets

Preliminary



Efficiency roughly independent of R used in clustering

Bottom line:

traditional and substructure techniques can be used together

[an analogous method still needs to be worked out for top]

Quark–gluon discrimination

It's potentially useful because
many signals are quark-rich,
while LHC backgrounds are gluon-rich

Quark and gluon jets are not trivial to define
[cf. Banfi, Zanderighi & GPS '06]
but we'll leave that aside for today

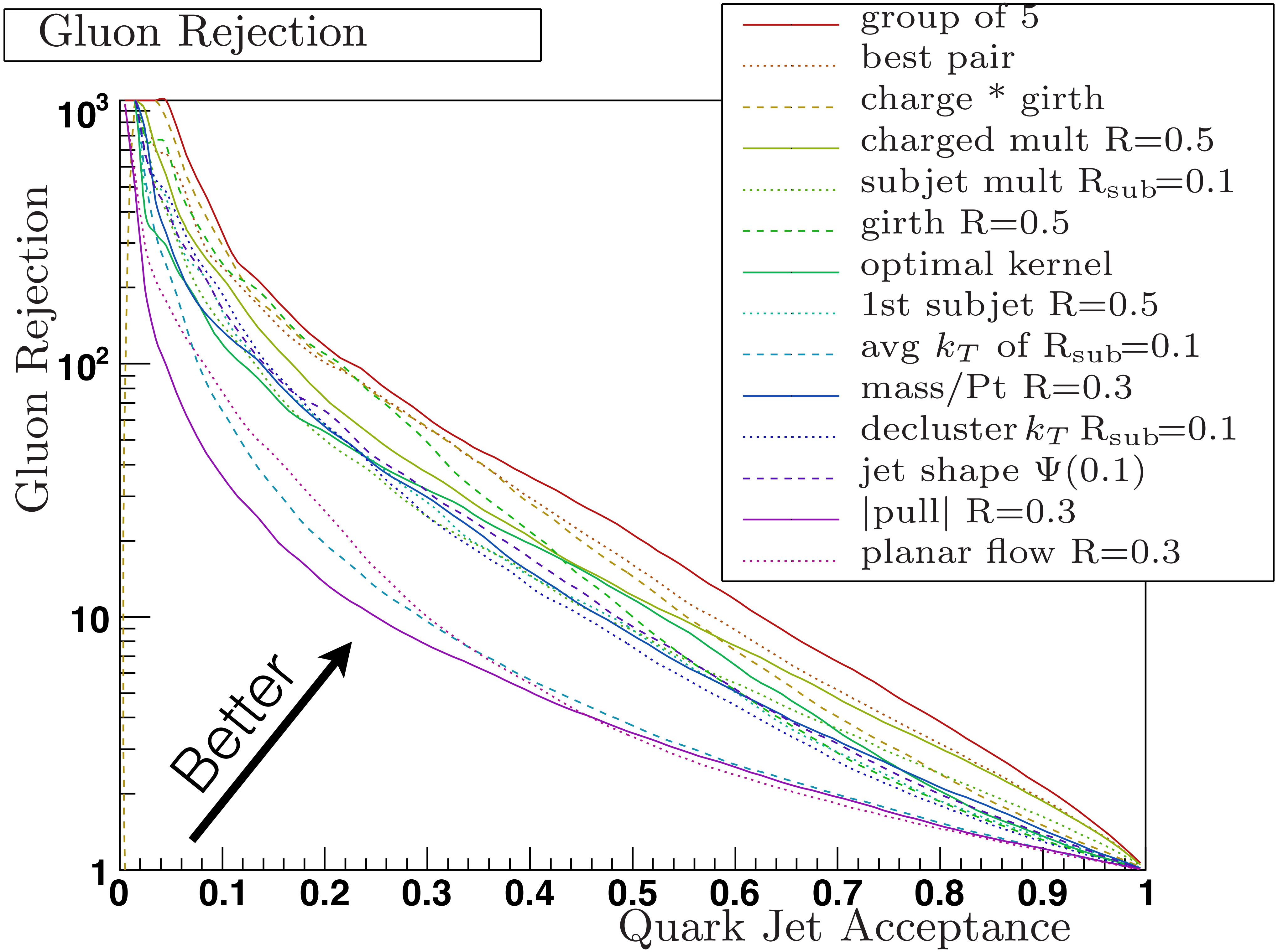
$$\boxed{\begin{array}{l}\text{Quarks radiate} \sim C_F = 4/3 \\ \text{Gluons radiate} \sim C_A = 3\end{array}}$$

So you want to study observables sensitive to the radiation inside the jet.

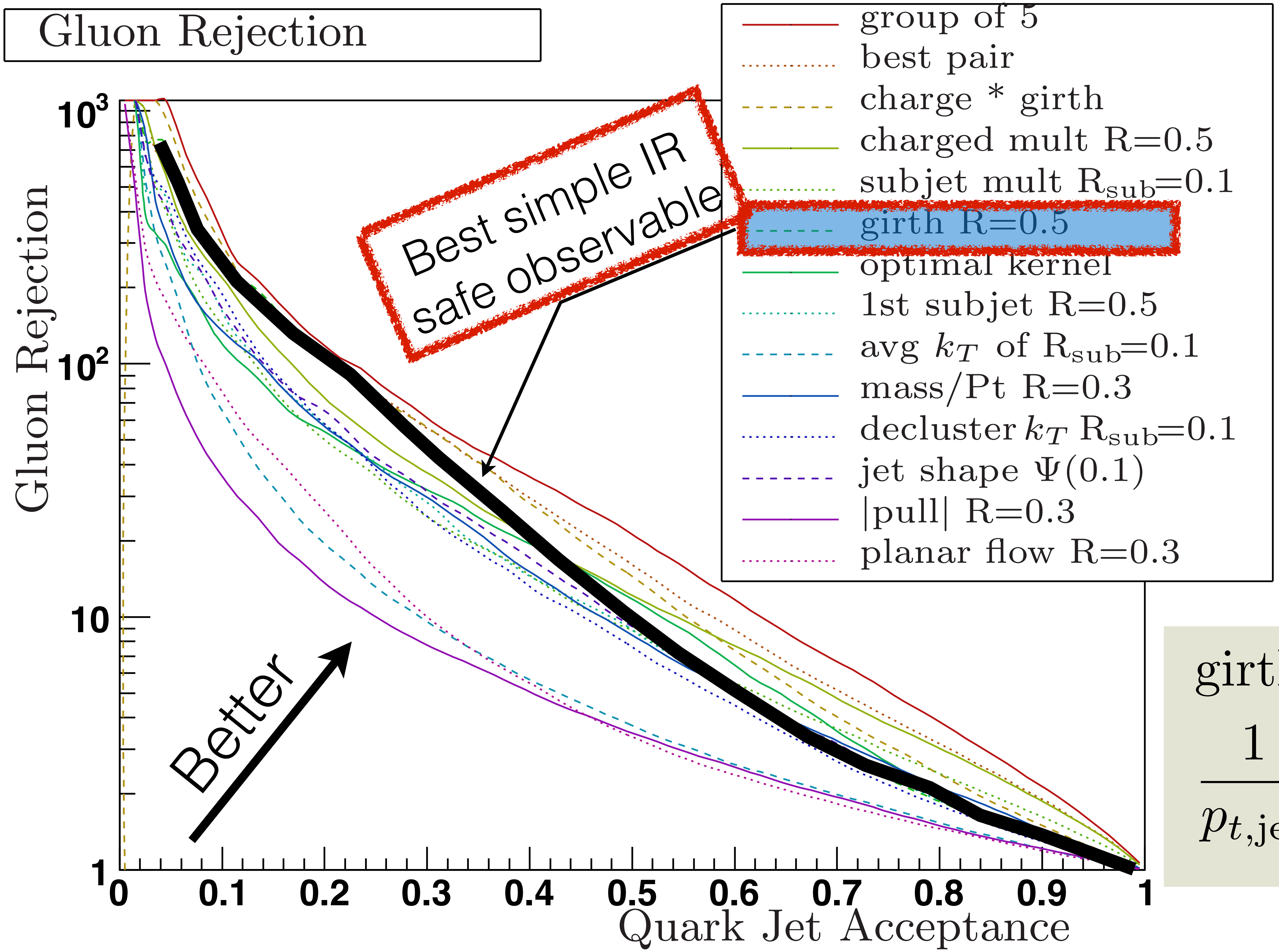
E.g. jet mass, particle multiplicity in jet, etc.

It's an old problem, but recently
Jason Gallicchio & Matt Schwartz
have been trying to attack it systematically

Gallicchio & Schwartz 1106.3076



Gallicchio & Schwartz 1106.3076



$$\text{girth} \equiv \frac{1}{p_{t,\text{jet}}} \sum_{i \in \text{jet}} p_{t,i}\, \Delta R_{i,\text{jet}}$$

$$A^{(1)} \equiv \text{girth} \equiv \text{broadening} \equiv \text{width} \equiv \frac{1}{p_{t,\text{jet}}} \sum_{i \in \text{jet}} p_{t,i}\, \Delta R_{i,\text{jet}}$$

What's the quantitative basis for quark–gluon discrimination?

[results in next pages from **Larkoski**, GPS & **Thaler**, in preparation]

$\Sigma_q(A^{(1)})$ = probability that quark jet has girth < $A^{(1)}$

$\Sigma_g(A^{(1)})$ = probability that gluon jet has girth < $A^{(1)}$

At leading logarithmic accuracy

$$\begin{cases} \Sigma_q(A^{(1)}) \simeq \exp\left[-C_F \frac{\alpha_s}{\pi} \ln^2 A^{(1)}\right] \\ \Sigma_g(A^{(1)}) \simeq \exp\left[-C_A \frac{\alpha_s}{\pi} \ln^2 A^{(1)}\right] = \boxed{\left[\Sigma_q(A^{(1)})\right]^{\frac{C_A}{C_F}}} \end{cases}$$

$$A^{(\beta)} \equiv \text{angularity}^{(\beta)} \equiv \tau_1^{(\beta,\text{jet-axis})} \equiv \frac{1}{p_{t,\text{jet}}} \sum_{i\in\text{jet}} p_{t,i}\, \Delta R_{i,\text{jet}}^{\beta}$$

At leading logarithmic accuracy

$$\begin{cases} \Sigma_q(A^{(\beta)}) \simeq \exp\left[-C_F \dfrac{\alpha_s}{\pi\beta} \ln^2 A^{(\beta)}\right] \\ \Sigma_g(A^{(\beta)}) \simeq \exp\left[-C_A \dfrac{\alpha_s}{\pi\beta} \ln^2 A^{(\beta)}\right] = \boxed{\left[\Sigma_q(A^{(\beta)})\right]^{\frac{C_A}{C_F}}} \end{cases}$$

Quark-gluon efficiency relation is independent of β at LL!

$$A^{(\beta)} \equiv \text{angularity}^{(\beta)} \equiv \tau_1^{(\beta,\text{jet-axis})} \equiv \frac{1}{p_{t,\text{jet}}} \sum_{i\in\text{jet}} p_{t,i}\, \Delta R_{i,\text{jet}}^{\beta}$$

At leading logarithmic accuracy

$$\begin{cases} \Sigma_q(A^{(\beta)}) \simeq \exp\left[-C_F \frac{\alpha_s}{\pi\beta} \ln^2 A^{(\beta)}\right] \\ \Sigma_g(A^{(\beta)}) \simeq \exp\left[-C_A \frac{\alpha_s}{\pi\beta} \ln^2 A^{(\beta)}\right] = \left[\Sigma_q(A^{(\beta)})\right]^{\frac{C_A}{C_F}} \end{cases}$$

Quark-gluon efficiency relation is independent of β at LL!

At NLL accuracy

$$\ln \Sigma_g = \frac{C_A}{C_F}\left(1 + \frac{4\pi}{12}\frac{C_A - C_F}{\beta}\alpha_s\right)\ln \Sigma_q$$

based on formulas from CAESAR

at NLL, smaller β gives better discrimination [but formula valid only for β > 1, saturates/modified for β ≤ 1]

$$C_1^{(\beta)} \equiv \text{energy-energy-correlation moments} \equiv \frac{1}{p_{t,\text{jet}}^2} \sum_{i,j \in \text{jet}} p_{t,i} p_{t,j} \, \Delta R_{i,j}^{\beta}$$

related observables had been studied analytically in CAESAR paper
& in past couple of years by Jankowiak & Larkoski
& in parallel with our work by Gallicchio & Schwartz in MC

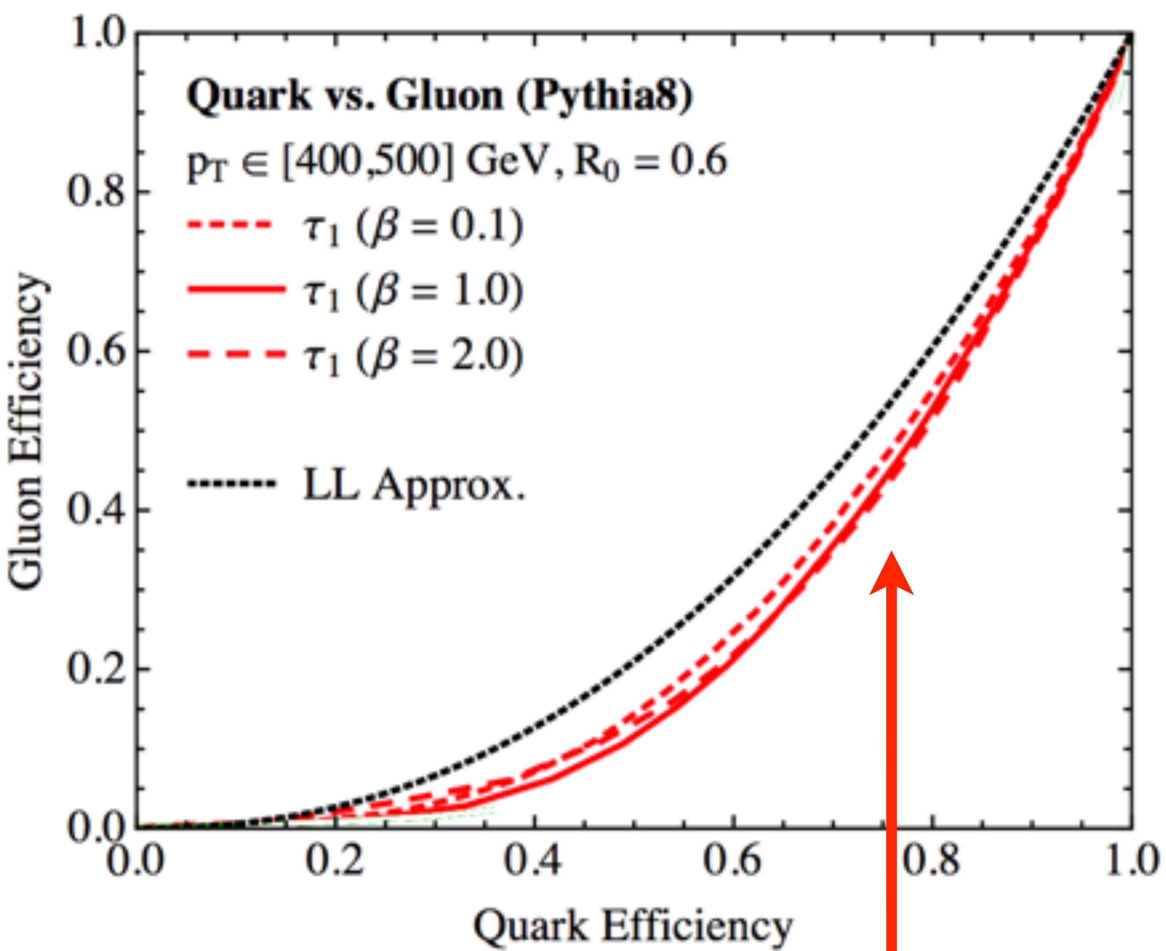
At NLL accuracy

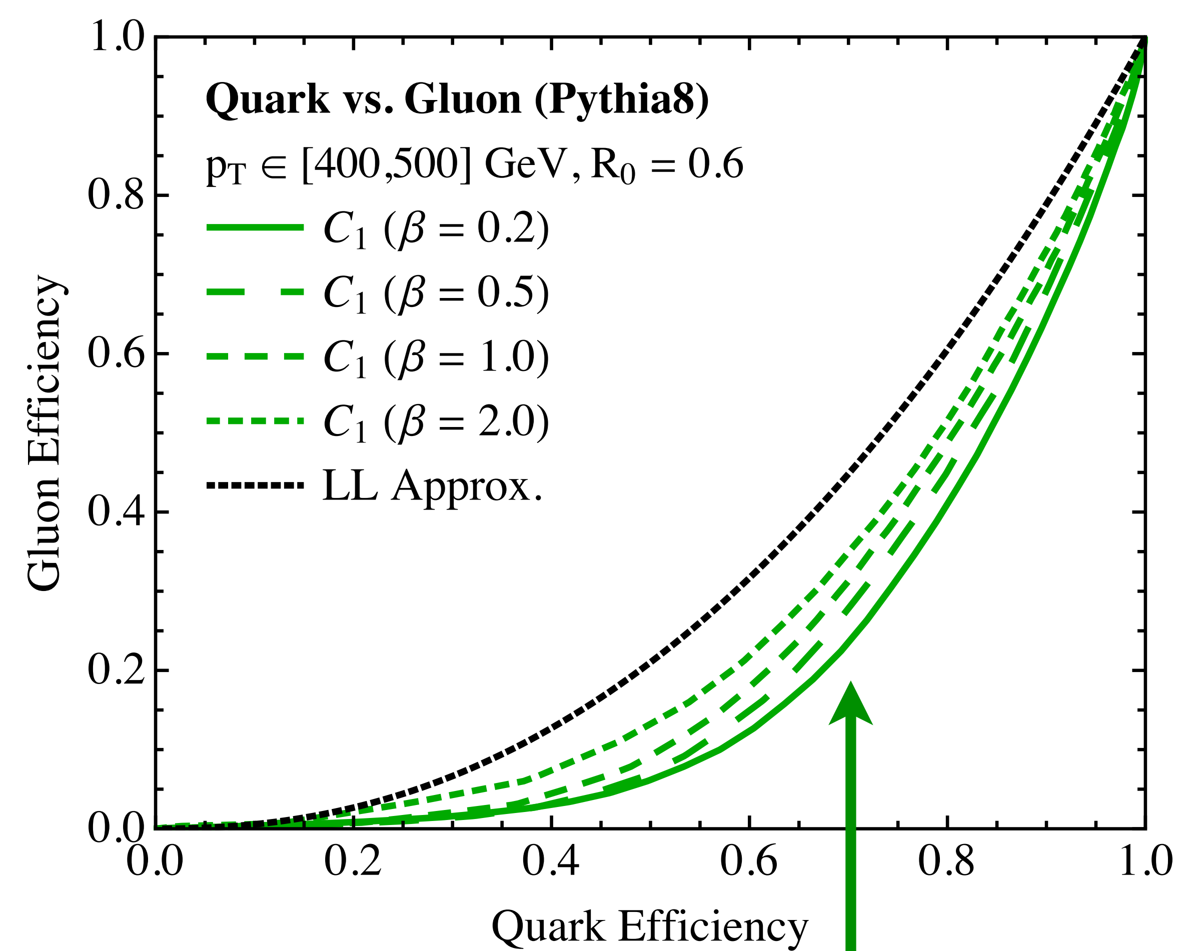
$$\ln \Sigma_g = \frac{C_A}{C_F} \left(1 + \frac{4\pi}{12} \frac{C_A - C_F}{\beta} \alpha_s \right) \ln \Sigma_q$$

same formula as before

at NLL, smaller β gives better discrimination

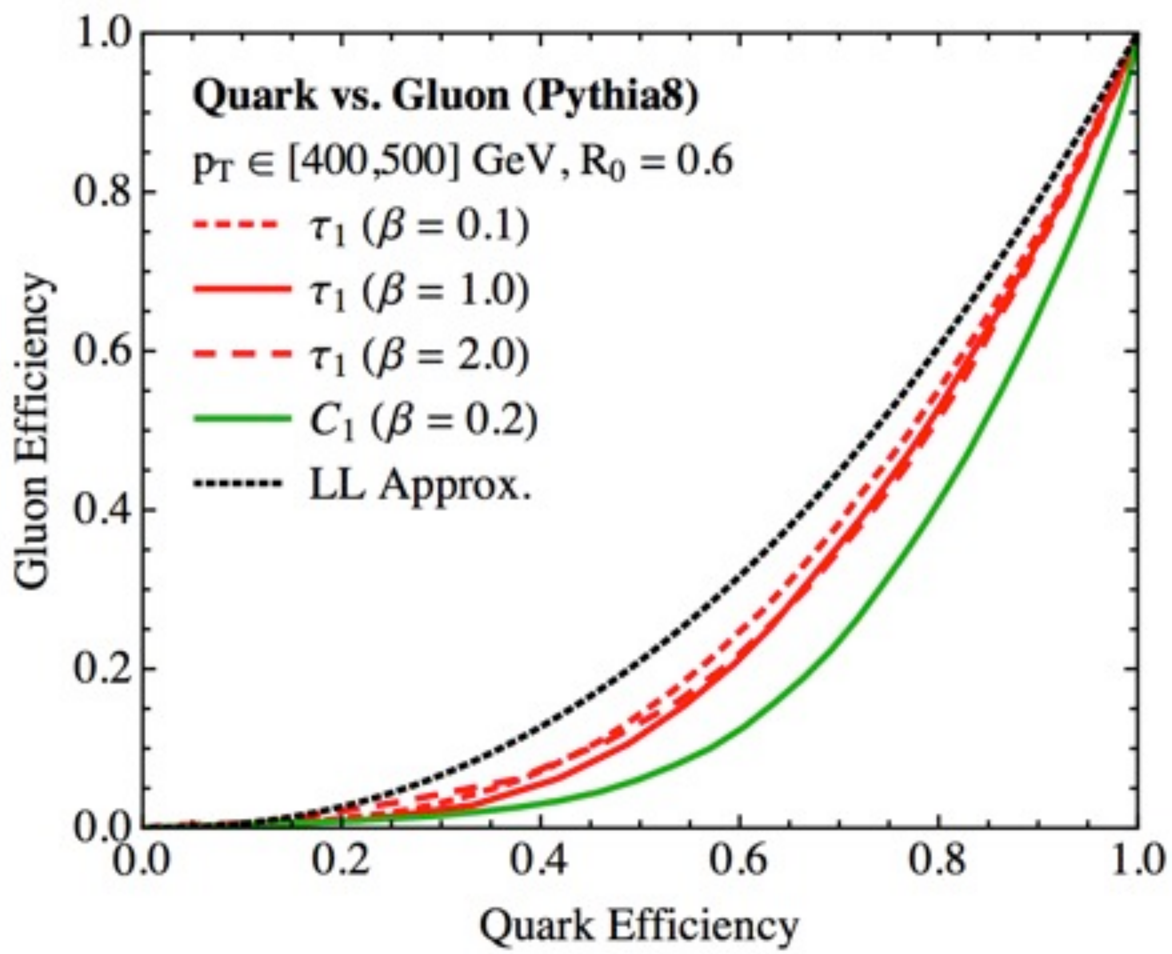
BUT: formula now valid for β > 0 rather than β > 1

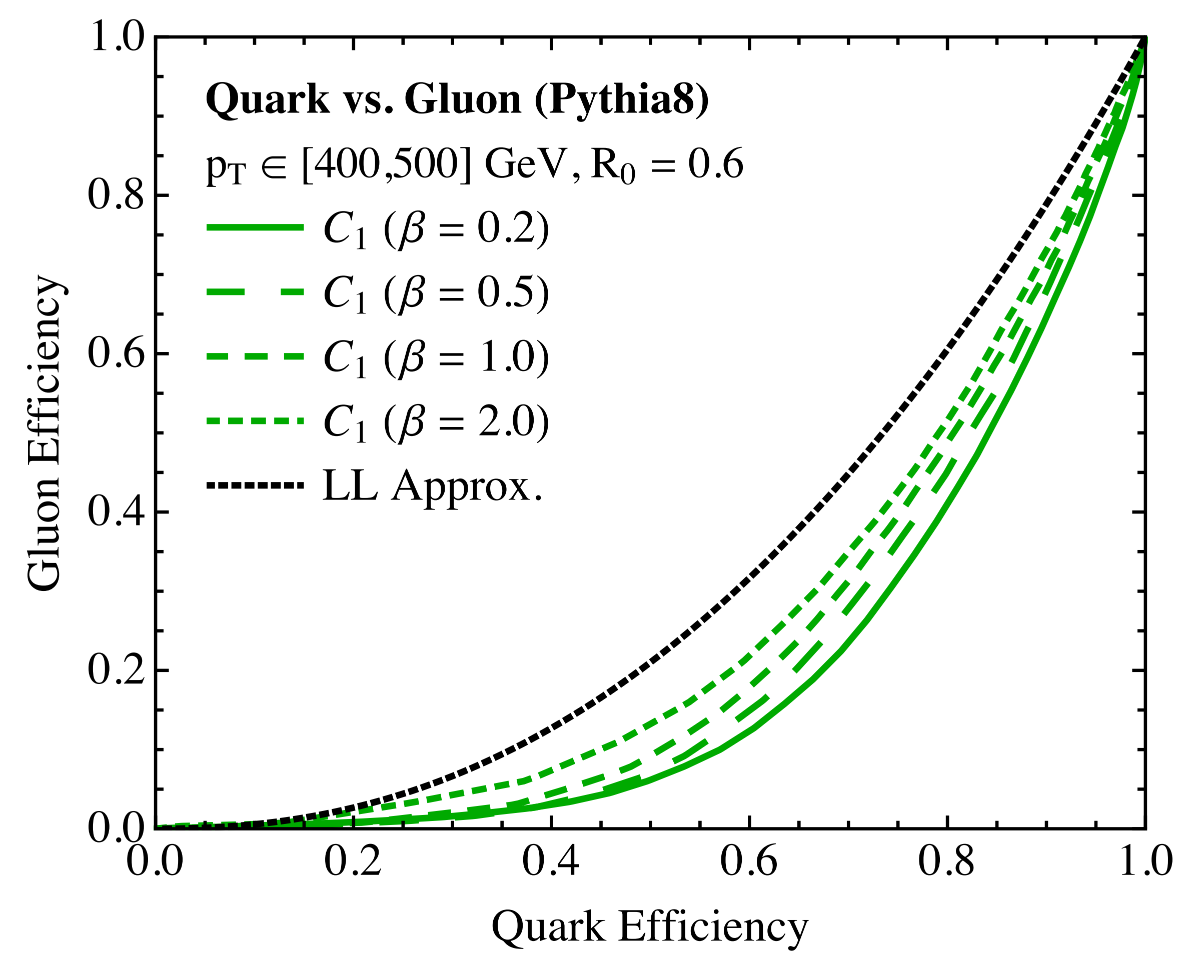




angularities saturate at β = 1

EEC moments continue to improve for β < 1





analytical understanding points to using
energy-energy correlation moments
which bring up to a factor ~2 better gluon rejection
than the jet width

[NB: less improvement with Herwig++]

Other things I would have liked to talk about

Q-jets

Does clustering have to give a unique answer? What if you probe multiple possible clustering histories?

Ellis et al. arXiv:1201.1914

Jet substructure by accident

Rather than looking for 16 jets (e.g. in BSM → 4 low-p_t tops), look for O(4) fat jets, each with substructure. May be easier to reliably predict backgrounds

Cohen, Izaguirre, Lisanti & Lou arXiv:1212.1456

Pileup

How to remove it from jet shapes

Cacciari et al. arXiv:1211.2811

Other things I would have liked to talk about

FastJet Contrib
A space for people to contribute their own jet-tool libraries, to provide users with uniform, regularly updated and reliable access to a broad range of jet tools.

fastjet.hepforge.org/contrib/

FastJet Contrib

The fastjet-contrib space is intended to provide a common location for access to 3rd party extensions of FastJet.

Download the current version: fjcontrib-1.001 (released 23 February 2013), which contains these contributions.
Changes relative to earlier versions are briefly described in the NEWS file.

After downloading and unpacking, enter the fjcontrib-1.001/ directory and then run

```
./configure [--fastjet-config=FILE] [--prefix=...] [...]
make
make check          # optional
make install
```

By default the package installs to the same directories as the FastJet installation.

A contribution named "SomeContrib" is usually accessed by including "fastjet/contrib/SomeContrib.hh" in your C++ file, and linking with -lSomeContrib.

Summary

Use of jets beyond the *"jet=parton"* idea is with us today.

That puts a responsibility on theorists to start understanding jet substructure beyond simply running Monte Carlos.

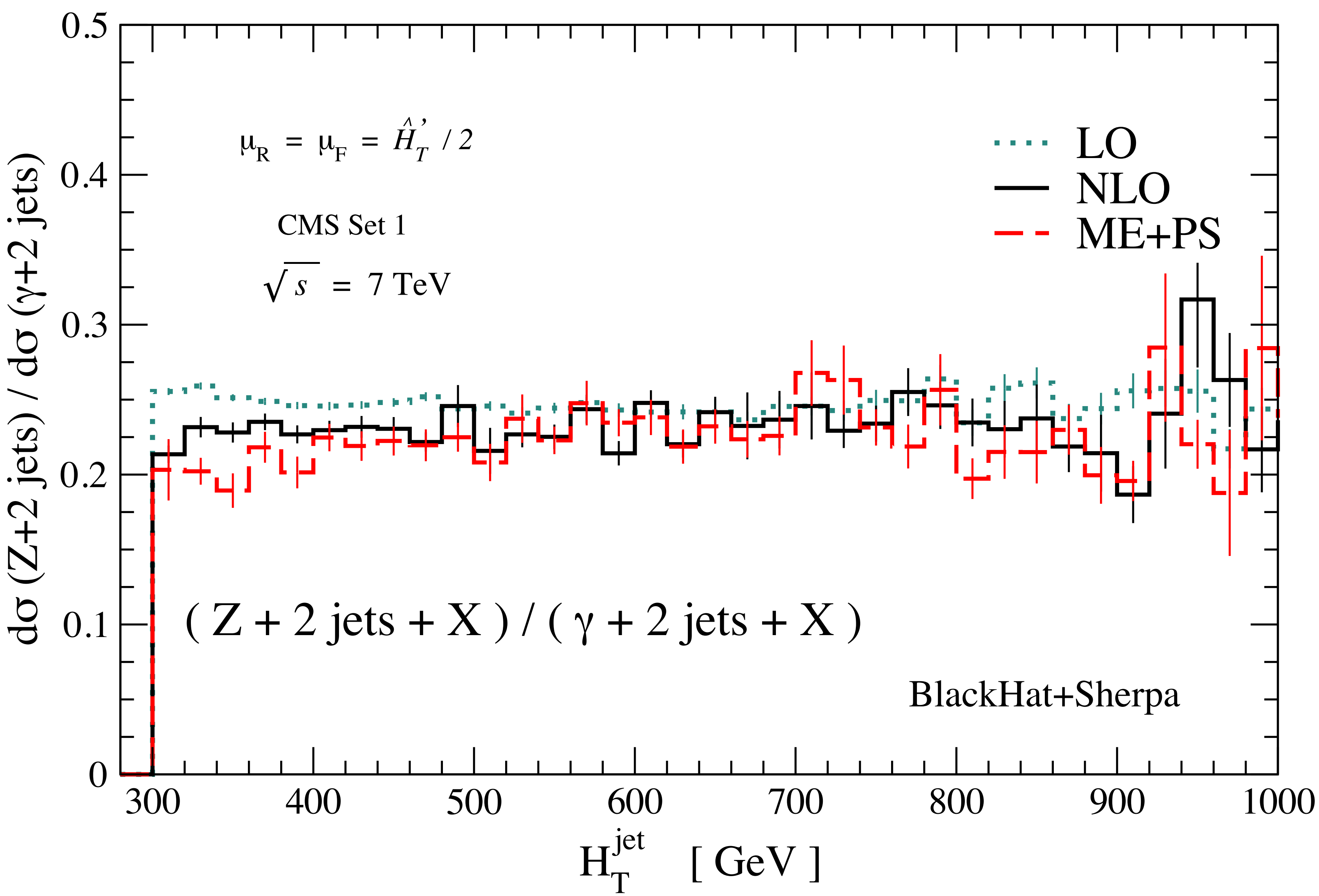
It seems that's feasible, with the potential also to guide development of more powerful and more robust jet tools.

Hopefully, this will help reliably stretch the boundaries of what LHC can do in its searches and measurements!

EXTRAS

Understanding your taggers means you know what tools you can safely use with them

For robustness, you can then choose taggers whose distributions can be predicted in many ways



Just like MET(Z→νν) in multijets is reliably estimated from γ+jets because multiple types of calculations of the ratio agree

What are we comfortable with?

Resolved Analysis

Find one jet/prong

Cut on jet p_t, Δy, ...

What are we comfortable with?

Resolved Analysis		Fat-jet Analysis
Find one jet/prong	→	Find subjets
Cut on jet p_t, Δy, ...	→	Cut on subjet z, ΔR, ...

[MDT/Prune/Trim/Filt/XYZTopTagger/Template ...]

What are we comfortable with?

Resolved Analysis		Fat-jet Analysis
Find one jet/prong	→	Find subjets
Cut on jet p_t, Δy, ...	→	Cut on subjet z, ΔR, ...
		[MDT/Prune/Trim/Filt/XYZTopTagger/Template ...]
Isolation cut for colourless leptons, γ		
Cut on radiation in jet for q/g discrimination		

What are we comfortable with?

Resolved Analysis		**Fat-jet Analysis**
Find one jet/prong	→	Find subjets
Cut on jet p_t, Δy, ...	→	Cut on subjet z, ΔR, ... [MDT/Prune/Trim/Filt/XYZTopTagger/Template ...]
Isolation cut for colourless leptons, γ	→	Cut on radiation for colourless W,H,...
Cut on radiation in jet for q/g discrimination	→	Cut on radiation in subjets [τ_{mn}, Qjets, deconstruction...]

What are we comfortable with?

Resolved Analysis		Fat-jet Analysis
Find one jet/prong	→	Find subjets
Cut on jet p_t, Δy, ...	→	Cut on subjet z, ΔR, ...
		[MDT/Prune/Trim/Filt/XYZTopTagger/Template ...]
Isolation cut for colourless leptons, γ	→	Cut on radiation for colourless W,H,...
Cut on radiation in jet for q/g discrimination	→	Cut on radiation in subjets
		[τ_{mn}, Qjets, deconstruction...]

Standard, well understood

Less standard requires case-by-base validation

Different fat-jet tagger types

Prong based

(e.g. HEPTopTagger, Template Tagger)

- Identifies prongs
- Requires prongs be consistent with kinematics of $t \to Wb \to$ 3 quarks

Radiation based

(e.g. N-subjettiness = τ_3/τ_2 + mass cut)

- Requires top-mass consistency (maybe with some grooming)
- Exploits weaker radiation from top (3 quarks) than background (1q+2g or 3g)

Infrared safety

Infrared safety:

When the addition of one soft particle with momentum **ε** changes the outcome of tagging by an amount O(1).

It means that perturbative calculations give ∞

It means that the physics of hard objects may be irremediably contaminated by non-perturbative physics – not good for robustness!

Was long an issue in hadron-collider jet-finding.
Let's make sure it doesn't come back to haunt us!

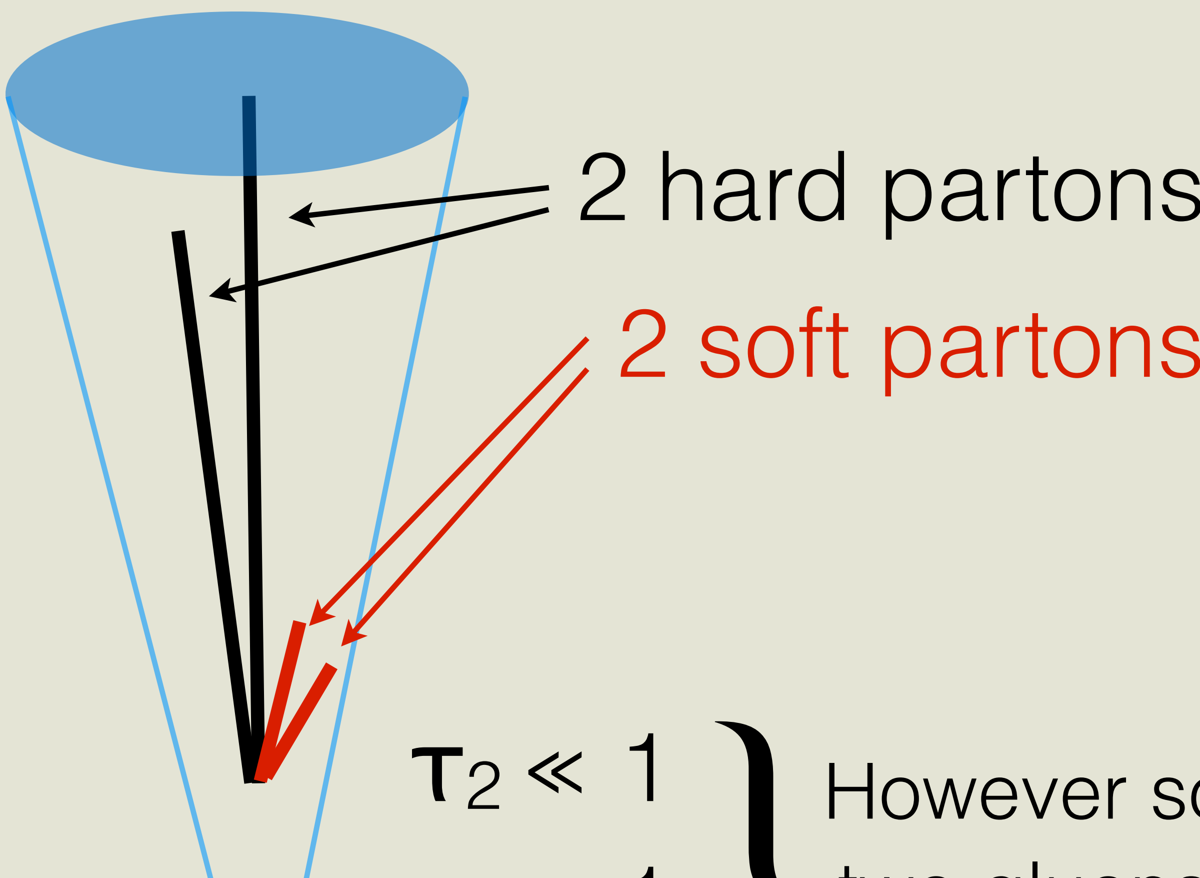
CMS's pruning followed by a mass-drop cut:

see blackboard!

IR issues in τ_{23}

N-subjettiness τ_3 / τ_2:

τ_2 measures departure from 2-parton energy flow
τ_3 measures departure from 3-parton energy flow



$$\left.\begin{array}{r} \tau_2 \ll 1 \\ \tau_3 \ll 1 \\ \tau_3/\tau_2 \sim 1 \end{array}\right\}$$

However soft the two gluons, you get finite τ_3/τ_2

That's IR unsafe

Easily cured with a cut on τ_2 / τ_1, which forces 3rd prong not to be soft.

Extra cut has almost no impact on performance

Cacciari et al '12

Pileup in the boosted regime

Pronged top taggers

Some have pileup-***reduction*** built in (HEPTopTagger, Template), essentially by using small (R~0.2–0.3) sub-cones, sometimes dynamically adjusted to the top p_t

For heavy pileup you will need to supplement them with full pileup ***subtraction*** (e.g. area-based).

[Technically trivial, but so far studied only for filtering & trimming]

Shape-based taggers

Until recently, no clear way of subtracting pileup.

Pileup subtraction for shapes

Cacciari, Dutta, JH Kim, GPS & Soyez '12

n^{th} **derivative of shape** wrt ghost momenta

Shape as a function of particle momenta in jet

$$V_{\text{jet}}^{[n]} \equiv A_g^n \frac{d^n}{dr_{t,g}^n} V(\{p_i\}_{\text{jet}})$$

Ghost area

Pileup subtraction for shapes

Cacciari, Dutta, JH Kim, GPS & Soyez '12

n$^{\text{th}}$ **derivative of shape** wrt ghost momenta

Shape as a function of particle momenta in jet

$$V_{\text{jet}}^{[n]} \equiv A_g^n \frac{d^n}{dr_{t,g}^n} V(\{p_i\}_{\text{jet}})$$

Ghost area

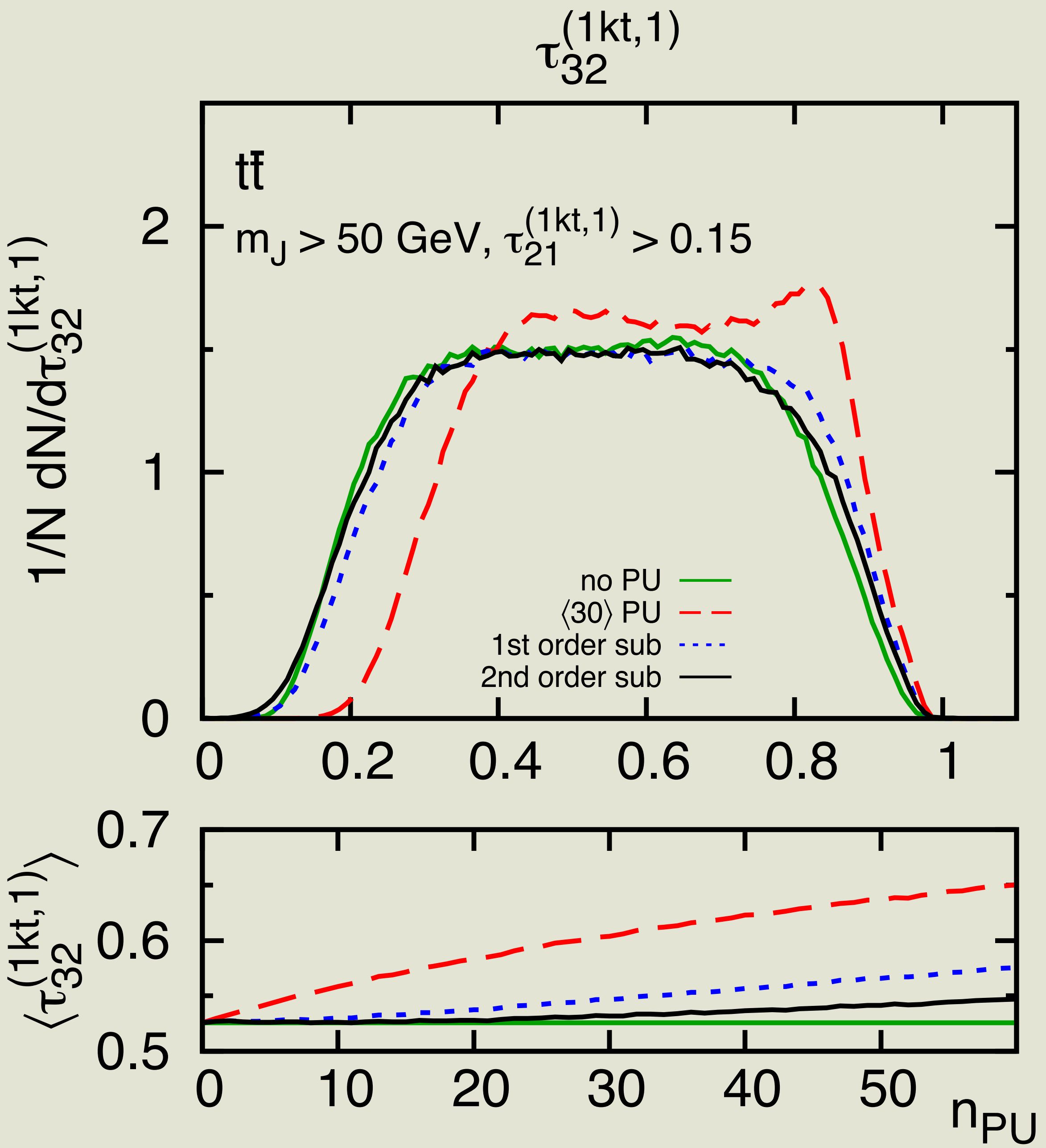
$$V_{\text{jet,sub}} = V_{\text{jet}} - \rho V_{\text{jet}}^{[1]} + \frac{1}{2}\rho^2 V_{\text{jet}}^{[2]} + \cdots$$

Subtracted shape

pileup density

Practical test: τ_{32} and top tagging

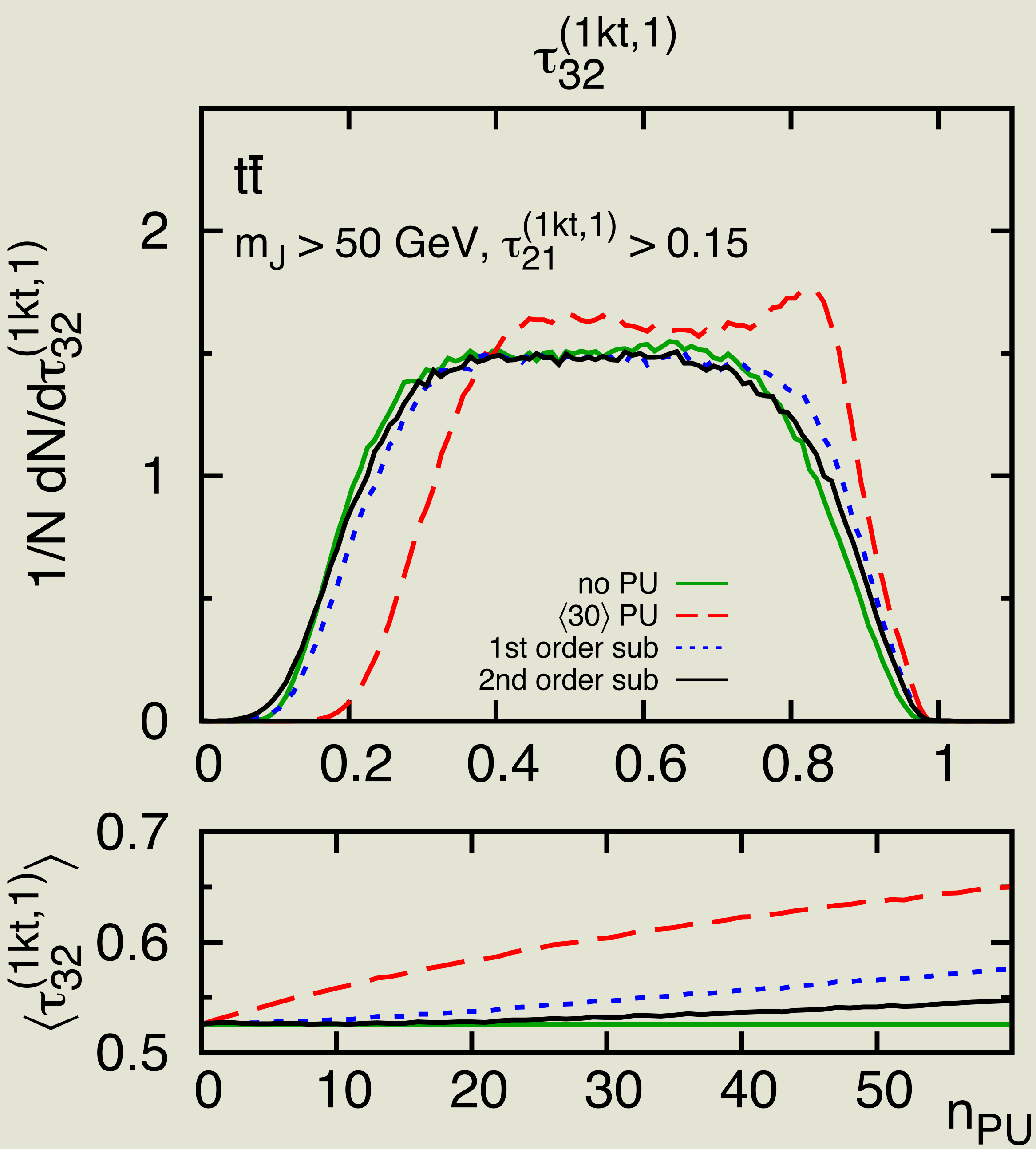
Correcting the τ_{32} distribution



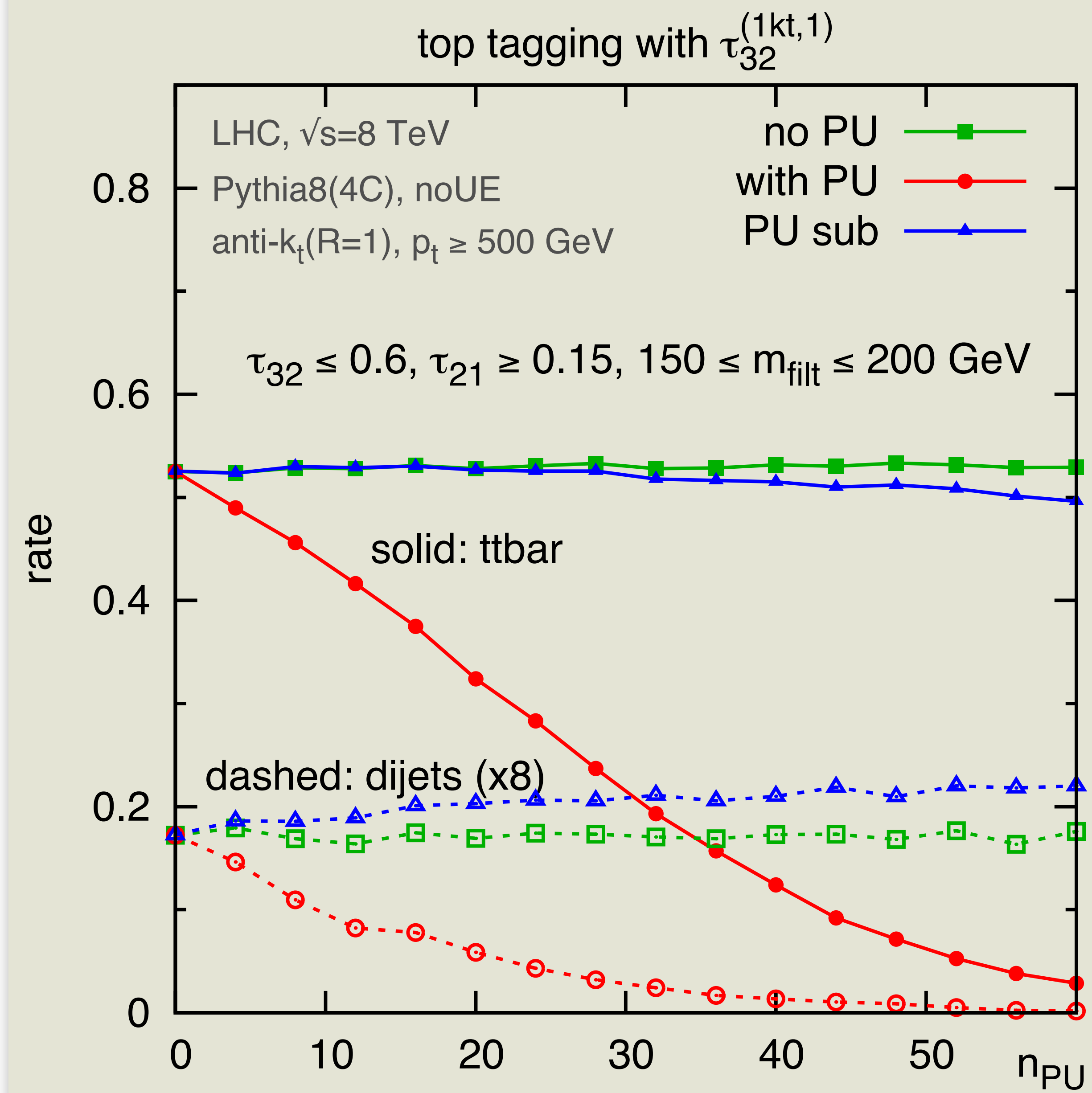
Green: no PU **Red: with PU** **Blue/Black: subtracted**

Practical test: τ32 and top tagging

Correcting the τ32 distribution



Tagging efficiency



Green: no PU **Red: with PU** **Blue/Black: subtracted**